Transcriptomic Approaches to Understanding Ageing and Metabolism in Rodents

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy

By

Andrew Peter Holmes

December 2014

Abstract

Ageing is a major cause of diseases in modern society and leads to many age-related diseases. In addition, rising obesity in society increases the likelihood of various diseases. The development of genetic sequencing technologies provides a novel platform to unravel the complexity of ageing and metabolism.

The Naked mole-rat (*Heterocephalus glaber*) is a rodent that is related to mice (*Mus musculus*) and ages very slowly, and can live for over 30 years without developing cancer. RNAseq analysis of *H. glaber* and *M. musculus* livers identified genes that were under- or over-expressed in *H. glaber*. Relative qPCR analysis was performed to confirm the expression of highly expressed genes in the naked mole-rat.

Adenosine-to-Inosine (A-to-I) RNA editing is a post-transcriptional modification of specific bases that can cause an alteration in amino acid and splicing events. Recent studies indicated a gene-specific decline in RNA editing with age in humans in *Cyfip2*. Using SOLiD RNAseq, we sequenced the transcriptomes of 6-, 12-, and 28-month old rat cerebral cortex. We identify a conserved RNA editing site in *Cog3.* Upon analysis of known conserved RNA editing targets, no changes in RNA editing were observed during ageing in the rat. These results highlight the biological differences between rodent model organisms and humans, and their significance in the context of RNA editing and ageing.

Dietary restriction of caloric intake is known to increase the lifespan of many species. Physiologically, these animals remain lean and show a later onset of age-related diseases. Through this concept, we were interested in studying Gnasxl$^{m+/p-}$ knockout mice, which remain lean due to hypermetabolism, despite increased food intake. Gnasxl$^{m+/p-}$ mice have a deletion in *Gnas,* which prevents expression of Xlαs. However, it is not known exactly how the loss of Xlαs exerts this phenotype. Due to its role in homeostasis and localisation of Xlαs expression, RNA from whole hypothalami of wildtype and Gnasxl$^{m+/p-}$ mice were analysed by Illumina RNAseq in order to identify expression changes that may help to explain the physiological symptoms. The glial cell marker *Gfap* was found to be downregulated 2-fold in Gnasxl$^{m+/p-}$ mice, which was confirmed by qPCR. Furthermore, in analysing the localisation of GFAP within the hypothalamus, we identified a distinct loss of GFAP expression in the arcuate nucleus, and suprachiasmatic nucleus. GFAP was not found to be decreased in postnatal (P)1 and P5 mice, indicating that the decrease in GFAP is more likely to be an adaptation to chronic undernutrition as a result of hypermetabolism, than a developmental problem in the mice.

Acknowledgements

I would like to thank my supervisors, Dr. João Pedro de Magalhães and Dr. Patricia Murray for their advice and guidance during this project. I would also like to thank Dr. Antonius Plagge for his support while welcoming me into his laboratory group, and for sharing his expertise.

I would also like to thank the many friends and family who have supported me over the years. To my mum for her unwavering support, and to the many friends who have been a shoulder to lean on during the hard times.

Finally, I would also like to thank the Biotechnology and Biological Sciences Research Council for financially supporting this project.

Contents

List of Figures

List of Tables

List of abbreviations

3V – Third ventricle

α-MSH – Alpha-melanocyte-stimulating hormone

A-to-I – Adenosine to Inosine

AC – Adenylyl cyclase

Actb – Beta actin

ACTH – Adenocorticotropic hormone

ADARs – Adenosine deaminases that act on RNA

AHO – Albright's hereditary osteodystrophy

AgRP – Agouti-related peptide

Arc – Arcuate nucleus of the hypothalamus

AT1R - Angiotensin II receptor type 1 receptor

ATP – Adenosine triphosphate

BMP – Bone morphogenetic protein

bp – Base pairs

cAMP – Cyclic adenosine monophosphate

cDNA – Complementary deoxyribonucleic acid

CNS – Central nervous system

CRH – Corticotropin-releasing hormone

CRHR - Corticotropin-releasing hormone receptor

DAPI - 4',6'-diamidino-2-phenylindole

ddPCR – Digital Droplet polymerase chain reaction

DMH – Dorsomedial nucleus

DNA – Deoxyribonucleic acid

ds – Double stranded

E – Embryonic day

ECL – Enhanced chemiluminescence

EST – Expressed sequence tag

FPKM - Fragments per Kilobase of exon per Million fragments mapped

G6P – Glucose-6-phosphate

Gαs – G protein α subunit

GABA – Gamma-aminobutyric acid

Gabarap – Gamma-aminobutyric acid receptor-associated protein

GDP – Guanosine diphosphate

Gfap – Glial fibrillary acidic protein

Gnasxl$^{m+/p-}$ — Mouse line with a defective paternal copy of the *Gnasxl* exon

GnRH – Gonadotropin-releasing hormone

GO – Gene ontology

GPCR – G protein-coupled receptor

GTP – Guanosine triphosphate

H6PD – Hexose-6-phosphate dehydrogenase

HA – High molecular mass hyaluronic acid

Hadh – Hydroxyacyl-CoA dehydrogenase

HPA Axis – Hypothalamus-Pituitary-Adrenal axis

HRP – Horseradish peroxidase

HSD11B1 – Hydroxysteroid 11-beta dehydrogenase 1

LH – Lateral hypothalamus

LNA – Locked nucleic acid

MC4R – Melanocortin-4-receptor

ME – Median eminence

mRNA – Messenger Ribonucleic acid

NADH – Nicotinamide adenine dinucleotide

NADPH - Nicotinamide adenine dinucleotide phosphate

NMR – Naked mole-rat

NPY – Neuropeptide Y

P – Postnatal day

PAGE – Polyacrylamide gel electrophoresis

PBS – Phosphate buffered saline

PCR – Polymerase Chain Reaction

PFA - Paraformaldehyde

PHP – Pseudohypoparathyroidism

Ppia – Cyclophilin A

PseudoPHP – Pseudopseudohypoparathyroidism

POMC - Proopiomelanocortin

PTH – Parathyroid hormone

PVN – Paraventricular nucleus

qPCR – Quantitative Real-time reverse transcriptase polymerase chain reaction

RIN – Ribonucleic acid integrity number

RNA – Ribonucleic acid

RNAseq – RNA sequencing

ROS – Reactive oxygen species

rRNA – Ribosomal ribonucleic acid

SCN – Suprachiasmatic nucleus

SDS – Sodium dodecyl sulphate

SDS-PAGE – Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SHH – Sonic Hedgehog

SINE – Short interspersed repeat element

SGS – Second generation sequencing

SNP – Single nucleotide polymorphism

SNS – Sympathetic nervous system

VMH – Ventromedial nucleus

XLαs – Extra Large G protein α subunit

Chapter 1: Introduction

Two of the largest challenges in modern society are ageing populations, and increases in obesity levels. Ageing is a constant process through which many species decline in health and fertility and have an increase in mortality. Ageing is a huge risk factor in developing the most challenging chronic diseases, such as diabetes, heart disease, cancer, and dementia. Therefore, the aim of ageing research is ultimately to prevent age-related diseases and to increase the time spent in good health, not just lifespan. Like ageing, obesity is a risk factor for many diseases. Studying the mechanisms behind the regulation of metabolism may lead to treatments in the future. The explosion of DNA sequencing technologies over the last decade means that identifying whole transcriptome gene expression changes has become more cost-effective, and produces more data than ever before. We sought to use these sequencing technologies as platforms to interrogate aspects of ageing and metabolism in rodent models.

1.1 Ageing

Ageing is a process that manifests itself physiologically. Particularly with elderly people, this can mean a deterioration in health. However, understanding exactly how this process occurs at the cellular level is much more challenging.

Many theories have been proposed as a way to explain the process of ageing. These include theories based on accumulated damage, such as the free-radical theory. This idea suggests that during oxidative phosphorylation in the mitochondria, reactive oxygen species (ROS) are produced, which are able to react with and damage DNA and proteins which lead to the deterioration of cells. Other theories suggest that ageing is genetically engineered. These postulate that it is not damage accumulation that drives ageing, but that ageing is a developmental process that is genetically determined by our genes.

Calorie restriction is a process that has been studied in many species. It consists of chronic underfeeding, but without the deleterious effects of malnutrition. In many species, calorie restriction has been shown to slow the ageing process, enabling individuals to live up to 50% longer in some species. Furthermore, the effects are observed in a wide range of species from yeast to non-human primates. Calorie restriction induces changes in metabolism, particularly in favouring metabolism of fat over that of carbohydrates. While multiple signalling pathways have been implicated in mediating the effects of calorie restriction, the

exact mechanism is not fully understood (Bordone and Guarente, 2005; Masoro, 2005; Speakman and Mitchell, 2011). Therefore calorie restriction has important implications for both ageing and metabolism – two of the key processes that are explored in this project.

In relation to the free-radical theory, the explanation is that consuming less food results in a lower production of ROS, leading to lower levels of DNA damage, which enables the cells to live longer.

Another important factor known to affect maximum longevity in mammals is body size (de Magalhães et al., 2007). Larger mammals, such as the African elephant can live for around 65 years; whereas smaller mammals such as the mouse live for a maximum of only 4 years (data from AnAge) (Tacutu et al., 2013). Some species do not entirely follow this pattern – a notable example being humans. While large compared to rodents, humans are much smaller than elephants – yet at 122 years, the maximum lifespan of humans exceeds that of elephants.

The increasing availability of genomic and transcriptomic data for many different species means that it is becoming easier than ever before to investigate not only the genetic basis of ageing, but in identifying similarities and differences between species. Most species show an age-related increase in mortality, whereby older individuals have a higher rate of mortality as a result of age-related diseases. However, there are some species that do not show signs of an age-related increase in mortality. One of these is the naked mole-rat (NMR), *Heterocephalus glaber*, which is particularly unusual in many aspects. Despite its body size, which is equivalent to that of a mouse, NMRs can live for over 30 years, and show a high resistance to developing cancer (Buffenstein, 2005; Tian et al., 2013). Therefore, like humans, NMRs have a much greater lifespan than would be expected for an animal of their body size. In addition, the NMR is a rare example of a eusocial mammal that has a queen within each colony. They live within burrows that contain a much lower concentration of oxygen than the atmosphere and maintain a very slow metabolic rate (Goldman et al., 1999).

As a result of it interesting longevity and cancer resistance, studying the physiology and genetics of the NMR has become an important focus for many aspects of both ageing and cancer research.

1.2 Metabolism

Many homeostatic mechanisms exist in species in order to regulate their responses to the environment. One key regulation is in energy management. After a period of not eating,

animals begin to feel hungry. Conversely, consuming a large meal will encourage satiety. Understanding the regulation of energy balance is of great importance considering the rising obesity levels in Western societies.

A key hormone involved in the regulation of feeding and appetite is leptin, which is produced from adipose tissue. A high amount of energy stored in the adipose tissue causes the release of leptin, which travels to the hypothalamus (Klok et al., 2007).The hypothalamus is a key region at the ventral side of the brain, which has crucial roles in homeostasis, circadian rhythm control, hormonal balances, and energy regulation. Leptin receptors on cells in the hypothalamus detect the levels of circulating leptin to activate neurons in the hypothalamus that act to suppress appetite. This forms a feedback cycle to prevent overeating.

In contrast, ghrelin is a hormone released by the gastrointestinal tract when the stomach is empty. Ghrelin travels to the hypothalamus to stimulate neurons in order to increase appetite. The hypothalamus acts as a mediator which responds to levels of these hormones in order to direct biological responses.

1.3 Second generation sequencing

The first papers on using second-generation sequencing (SGS) technologies to sequence transcriptomes were published in 2008 (Lister et al., 2008; Mortazavi et al., 2008; Nagalakshmi et al., 2008). Until this point, SGS had only been used for genome sequencing. However, alternative splicing events allow for multiple transcripts to be expressed from single genes. After post-transcriptional and post-translational modifications, the resulting proteins can have widely differing functions. Therefore the transcriptome is arguably much more informative than the genome in terms of understanding the biological functions of genes. Furthermore, with the rapidly increasing amount of genomic data on gained from SGS technologies, it became increasingly obvious that these genes needed to be annotated and characterised before we can understand how their transcripts function. Costs of SGS have reduced at a rapid rate since the introduction of SGS platforms, which has put them at the forefront of a revolution in genetic analysis.

RNAseq is a recently developed tool that employs SGS to sequence the transcriptome, as opposed to the genome. Rather than sequencing DNA, RNA is reverse transcribed to cDNA prior to sequencing. In sequencing RNA, it is possible to identify genes that are changed in expression between different conditions.

The main two SGS technologies at the time were SOLiD sequencing (Applied Biosystems) and Illumina sequencing. While ultimately they both produce output genomic and transcriptomic data, each method has its advantages and disadvantages.

1.3.1 SOLiD Sequencing

SOLiD sequencing uses a ligation-based method whereby 8-base probes bind to the cDNA library reads, which are attached to beads on the flow cell. The probes have four different fluorescent dyes attached to the final base. Which one is attached corresponds to the first two bases of the probe. After ligation, the final three bases are cleaved from the probe. This releases the dye, allowing the fluorescence to be detected by the machine. The next ligation begins at the sixth base, and the process continues for 7 cycles in total to create a sequence of 35 bases. The sequencing cycle is begun again after ligation of an initial probe to the second base rather than the first.

Because each colour refers to a 2-base sequence, each base is sequenced twice: once in combination with its previous partner, and once in combination with its next partner. Ultimately this results in one of the main benefits of SOLiD sequencing, which is its very high accuracy (99.94% for SOLiD v4) compared to other systems (Liu et al., 2012).

The result is that the read is identified not by the bases AGCT, but by a series of colours, which are converted to numbers 0-3. Depending on the identity of the first base, the colorspace read can correspond to four different sequences. Due to the length of the product however, during the mapping process, the correct sequence can be identified against the genome.

Therefore, the main benefit of the SOLiD sequencing system is for applications where base accuracy is very important. In particular, applications such as genome wide quantification of Single Nucleotide Polymorphisms (SNPs) would require a low error rate. However, applications such as gene expression changes are not so dependent on such a high accuracy. For these applications, long reads will still map to the correct gene even with a small number of sequencing errors.

1.3.2 Illumina Sequencing

Illumina sequencing is achieved using sequencing by synthesis. cDNA libraries are synthesised with adaptor sequences at either termini. These are adhered to the flow cell which contains a lawn of adaptor sequences. One phase of bridge amplification occurs, with the original cDNA washed away. This attaches the cDNA to the flow cell. cDNA is then amplified further by bridge amplification, which causes amplified cDNA to bend down so that

free distal cDNA adaptor sequence can adhere to a nearby free flow cell adaptor sequence. With several melting and amplification steps, this results in clusters of amplified cDNA originating from a single cDNA molecule. Dye-tagged nucleotides are flowed across, resulting in fluorescence given off by each cluster depending on which base is incorporated. The fluorescence wavelength is converted to the corresponding base, generating the digital reads.

Illumina sequencing has the main benefit of generating a huge amount of sequencing data. Illumina generates around 5 times as much genetic data as SOLiD (Liu et al., 2012). While accuracy is lower than SOLiD (98% compared to 99.94%) (Liu et al., 2012), the read length is longer at 100 base pairs (bp) compared to 35bp. The combination of these factors is that Illumina sequencing is able to generate data that has a wider genome coverage, and the massive amount of genomic data that it produces means that the cost per base is half that of the SOLiD system (Liu et al., 2012).

1.4 Aims

Building on previous work, identifying a decline in RNA editing with age in human cerebral cortex (Nicholas et al., 2010), we sought to use an RNA sequencing (RNAseq) approach to identify new conserved editing sites in the rat. Furthermore, we wanted to investigate whether editing of conserved RNA editing targets declines with age in the rat.

In addition, a recent RNAseq approach was used by colleagues to identify genes that were overexpressed in the liver of NMRs compared to their orthologous genes in mice – in order to identify genes that might be involved in the resistance to cancer and slow ageing.

RNAseq was employed to investigate an interesting phenotype in *Gnasxl*$^{m+/p-}$ mice. These mice have an exon deletion in the paternal copy of the imprinted gene *Gnas*, which encodes the G-protein α subunit. Interestingly, the mice are hypermetabolic and remain small and lean with decreased fat mass, despite consuming more food. Because of its role in energy balance, and a site of expression for the knocked-out XLαs, the hypothalamus was a key region of interest. In order to investigate genes that might be involved in the development of the phenotype, we aimed to use RNAseq to identify differentially expressed genes in the hypothalami of wildtype and *Gnasxl*$^{m+/p-}$ mice.

From this analysis, we identified several genes that were significantly differentially expressed. In order to verify these results, we chose to perform quantitative reverse transcriptase PCR (qPCR) on a subset of genes.

Upon verifying the expression of Glial fibrillary acidic protein (*Gfap*) as being underexpressed 2-fold in *Gnasxl*$^{m+/p-}$ mice, we wanted to analyse whether this was true of GFAP at the protein level. In order to do this, we employed the use of Western blots from whole hypothalamic lysates, and immunofluorescence on hypothalamic sections to analyse GFAP localisation. Furthermore, GFAP expression was analysed within the hypothalami of neonatal mice in order to determine whether changes in expression in the adult were due to problem during development of the hypothalamus, or whether it was as a result of chronic undernutrition.

Chapter 2: Adenosine to Inosine RNA editing does not change with age in the healthy male rat brain

2.1 Introduction

2.1.1 Adenosine-to-Inosine RNA editing

Adenosine-to-Inosine (A-to-I) RNA editing is a post-transcriptional process modulated by Adenosine Deaminases that Act on RNA (ADAR) proteins that convert specific adenosine residues to inosine by deamination. The inosine generated during editing is read as a guanosine by the splicing and translation machinery of the cell, as well as sequencing technologies. By inducing base modification, A-to-I RNA editing can have multiple effects on the splicing and sequence of mRNA and therefore efficacy of a protein. A-to-I RNA editing directs temporal-spatial regulation of transcripts and has an important role in increasing proteome diversity (Schaub and Keller, 2002; Nishikura, 2010).

2.1.1.1 RNA editing in health & disease

Many of the previously identified targets of RNA editing correspond to genes involved in neurotransmission in the central nervous system (CNS). In the CNS, RNA editing allows for fine regulation of the neurotransmission processes. Deregulation of A-to-I RNA editing is known to be a contributing factor in a variety of behavioural abnormalities and neurodegenerative diseases in the Central Nervous System (CNS), including Dyschromatosis Symmetrica Hereditaria (DSH), Amyotrophic Lateral Sclerosis (ALS), Alzheimer's disease and Huntington's disease, epilepsy, depression, schizophrenia and suicide (Akbarian et al., 1995; Maas et al., 2006; Farajollahi and Maas, 2010; Simmons et al., 2010). Furthermore, it has been hypothesised that RNA editing could be used as a therapeutic target for CNS diseases (Morabito and Emeson, 2009). The most widely studied example of RNA editing in the CNS is of the AMPA receptor subunit GRIA2. Editing in the coding region at the Q/R site is almost absolute and is required for tetramerisation of the subunits, which prevents Ca^{2+} ions from permeating the ion channel. Influxes of Ca^{2+} ions otherwise lead to excitotoxicity that results in damage and death of the motor neurons. *Adarb1*$^{-/-}$ mice die a few weeks after birth after suffering from epileptic seizures due to neuron death and excitotoxicity by Ca^{2+} influx (Brusa et al., 1995; Higuchi et al., 2000). ALS has also been linked to disruption of GRIA2 editing at the Q/R site, although it is not known whether this is causative.

In addition to having important roles in disease, ADARs appear to play a wide role in animal behaviour. Knockout fruit flies for *dADAR* (the sole ADAR homologue) show a selective disadvantage through reduced mating, accompanied by cerebral degeneration, erratic flight, uncoordinated locomotion and tremors. However, despite these abnormalities, *dADAR* mutants had a similar longevity to wild-type (Palladino et al., 2000; Jepson et al., 2010). A recent study showed that *ADARB1*$^{-/-}$*/Gria2*$^{R/R}$ mice, whereby the *Gria2* allele was modified to encode the edited isoform in the genome, also have behavioural abnormalities, including impaired hearing (Horsch et al., 2011). However, like the fruit fly, these mice have a normal lifespan. Microarrays on *ADARB1*$^{-/-}$*/Gria2*$^{R/R}$ mice revealed many genes altered in expression compared to wild type, hinting at downstream effects of a lack of ADARB1 and RNA editing. These studies demonstrate that ADARs have much wider roles than previously thought. Meanwhile, non-coding RNA species and various RNA processes are being increasingly linked to neurological and cognitive disorders in addition to ageing (Montano and Long, 2011; Qureshi and Mehler, 2011). Deregulation of RNA editing has long been associated with the incidence of various cancers, many of which beyond the CNS, indicating that disruption of RNA editing has far reaching effects in pathology (Klimek-Tomczak et al., 2006; Paz et al., 2007; Cenci et al., 2008; Martinez et al., 2008; Galeano et al., 2010). Dysregulation of RNA editing by overexpression of *Adarb1* results in obesity (Singh et al., 2007). In addition, *Adarb1* knockdown rat pancreatic cells showed disrupted exocytosis of, but not limited to, insulin in response to glucose (Yang et al., 2010).

Since deregulation of RNA editing may in part be responsible for a decline in neurological function, it was hypothesised that RNA editing may have important roles in neurological degeneration observed in ageing. A human centenarian study identified 5 SNPs in *Adarb1* and 13 SNPs in *Adarb2* (Bass et al., 1997) that were associated with extreme longevity in different genetic backgrounds (Sebastiani et al., 2009). This study raises further questions about the potential functions of ADARB2, although it could be an evolutionary relic. Though *Adarb2*$^{-/-}$ mice are said to be physiologically normal (Nishikura, 2010), the effect of the gene knockout across an organism's entire lifespan has yet to be studied. Recently it was demonstrated that RNA editing of *Cyfip2*, involved in synaptic maintenance, significantly declines with age, further strengthening the link between RNA editing and ageing (Nicholas et al., 2010). Editing of other targets including *Gabra3*, involved in neurotransmission; and the Q/R site in *Gria2*, did not significantly decline with age (Carlson et al., 2000; Nicholas et al., 2010). This indicates that age-related changes in RNA editing may be gene-specific.

2.1.1.2 Bioinformatics for RNA editing

Early studies demonstrated that bioinformatics could be exploited to identify new targets of A-to-I editing (Morse and Bass, 1999; Hoopengardner et al., 2003), although these studies only identified a handful of targets. A large-scale computational analysis later identified over 12000 potential target sites of A-to-I editing in over 1600 genes by analysis of mismatches between Expressed Sequence Tags (ESTs) and cDNA sequences (Levanon et al., 2004). A more recent genome-wide study in humans identified many potential A-to-I editing sites in exons outside Alu elements by identifying mismatches between genomic DNA and cDNA, demonstrating that there are still many uninvestigated targets of A-to-I editing (Li et al., 2009). Although RNA editing targets are much more prevalent in humans, many RNA editing targets are conserved between humans and rodents (Table 2.1).

Gene ID	Site	Genomic location	Reference
Kcna1		4:163012291	(Hoopengardner et al., 2003)
Cyfip2		10:31403167	(Levanon et al., 2005)
Gria2	R/G	2:172272247	(Lomeli et al., 1994)
	Q/R	2:172288193	(Higuchi et al., 1993)
Grik2	Q/R	20:53378633	(Barbon et al., 2003)
	Y/C	20:53408958	(Barbon et al., 2003)
	I/V	20:53408971	(Barbon et al., 2003)
Htr2c	A	X:31041849	(Burns et al., 1997)
	B	X:31041847	(Burns et al., 1997)
	E	X:31041843	(Niswender et al., 1998)
	C	X:31041842	(Burns et al., 1997)
	D	X:31041837	(Burns et al., 1997)
Gria3		X:3477800	(Lomeli et al., 1994)
Gabra3		X:158405626	(Ohlson et al., 2007)
Cog3		15:56613090	(Shah et al., 2009)
Flna		X:160382329	(Levanon et al., 2005)
Gria4		8:994881	(Lomeli et al., 1994)
Igfbp7		14:33010497	(Levanon et al., 2005)
Blcap	Y/C	3:148267368	(Clutterbuck et al., 2005)
	Q/R	3:148267359	(Clutterbuck et al., 2005)
	K/R	3:148267329	(Clutterbuck et al., 2005)
Adarb1	+10	20:11709472	(Rueter et al., 1999)
	+23	20:11709485	(Rueter et al., 1999)
	+24	20:11709486	(Rueter et al., 1999)
C1ql1	T/A	10:92116112	(Sie and Maas, 2009)
	102 Q/R	10:92116102	(Sie and Maas, 2009)
	093 Q/R	10:92116093	(Sie and Maas, 2009)
Grik1		11:27749588	(Barbon and Barlati, 2000)

Table 2.1: Conserved exonic editing sites in vertebrate species. A brief review of the known conserved targets of A-to-I RNA editing

Prior to splicing, repeat sequences known as Short Interspersed Repeat Elements (SINEs) are able to fold back on one another to form imperfect dsRNA regions. These SINEs are prevalent throughout mammals, as is the process of ADAR-mediated RNA editing. Since most A-to-I targets are in introns, sites can be targeted to either remove or create splicing sites. Perhaps the most interesting example is of ADARB1 itself, which can edit its own pre-mRNA. This creates an alternative proximal 3' acceptor site, leading to the inclusion of an extra 47bp in the mRNA after splicing. This results in the generation of a premature stop codon due to frameshift (Rueter et al., 1999). Self-regulation of ADARB1 allows control over editing levels (Feng et al., 2006). When in exons, editing of target sites may cause a direct alteration of the encoded base, altering protein functionality and increasing diversity. It has been hypothesised that editing may also alter efficiency of localisation to the cytoplasm and therefore expression (Maas et al., 2003).

Around 90% of A-to-I editing targets are found in Alu elements, a class of SINEs, which generate the dsRNA substrate for ADAR deamination upon fold-back (Morse and Bass, 1999; Kikuno et al., 2002; Kim et al., 2004; Levanon et al., 2004). The Alu elements account for 11% of the human genome and therefore have a high incidence of RNA editing targets (Deininger, 2011). Accordingly, it was identified that 85% of human pre-mRNAs were targets for A-to-I editing, whereas 1.4% of human mRNAs were edited (Athanasiadis et al., 2004), suggesting that the vast majority of editing events occur within introns, further supported by other analyses (Blow et al., 2004; Kim et al., 2004; Levanon et al., 2004). More recent studies have focussed on the use of Single Nucleotide Polymorphism (SNP) databases as a source of finding A-to-I editing targets. Many SNPs entered into this database were derived from their mRNA sequence alone. Therefore, many targets of A-to-I editing have been erroneously ascribed to be SNPs (Eisenberg et al., 2005a; Gommans et al., 2008). In addition, the NCBI Trace Archive, a repository for capillary and gel sequence data used in genome assembly, was recently scrutinised. As with the SNP databases, some sequences were RNA-derived. These sequence reads were aligned to the references genomes, identifying mismatches. After excluding sequencing errors, the authors discovered evidence for RNA editing, consistent with their distribution in SINEs and base preferences (Zaranek et al., 2010). Due to the rapid discovery of RNA editing targets from genome-wide studies, a new database was recently developed to hold this information, which will serve as a useful resource for researchers (Kiran and Baranov, 2010).

Though ADARs do not bind consensus sequences, they do display preferences for bases near to the target site. There is a strong preference against a G and for a U in the immediate 5' position to the target site, and a strong preference for a G in the immediate 3' position. Furthermore, on the sequence of the opposing side on the dsRNA structure, there is a preference for a C (Blow et al., 2004; Kim et al., 2004; Riedmann et al., 2008). A tool developed in 2011 is able to give predictions about potentially edited sites in a given sequence. Its algorithm, however, relies solely on preferences for the bases immediately 5' and 3' (Eggington et al., 2011). ADAR and ADARB1 both display slightly different preferences for these nearby bases (Lehmann and Bass, 2000; Riedmann et al., 2008). A more recent study has led to the discovery that editing of certain sites may be coupled to other nearby sites, spaced out at regular intervals. Therefore editing of a site may entirely depend on the presence or lack of editing at nearby sites (Ensterö et al., 2009). Importantly, ADARs are thought to bind to RNA sequences according to structure rather than sequence, demonstrating why prediction of editing sites is problematic.

Though RNA folding bioinformatic tools exist, they are each rather limited in terms of their accuracy (Mathews et al., 2010). The minimum free energy state represents the most physically stable structure that the RNA molecule can fold into. However, often this is not the biological state. Computationally, calculating RNA folding is complex due to the high numbers of potential interactions generated by even short sequences. Prediction of RNA editing sites by RNA structure would likely be extremely useful, but unfortunately this area needs much developing before it can be used in a large scale context.

2.2 Aims

Through a massive increase in genomic data, it has become possible to identify Single Nucleotide Polymorphisms (SNPs) at a genomic level through a process called SNP calling. This can be used for a wide variety of analyses, including identifying SNPs prevalent in cancerous tissues (Yu and Sun, 2013), or in the 1000 Genomes project, which seeks to identify human genetic variation (Abecasis et al., 2010). More recently, researchers have begun to use transcriptomic data and SNP calling as a way to identify and quantify RNA editing sites within the genome (Picardi et al., 2010; Park et al., 2012; Picardi et al., 2012; Ramaswami et al., 2013; Porath et al., 2014).

Previous research suggested a gene-specific decrease in RNA editing in the human cerebral cortex with age (Nicholas et al., 2010). Our aims were to sequence the transcriptomes of cerebral cortex from 6-, 12-, and 28-month old rats, using SOLiD v4 sequencing. This method

of SGS was chosen because of its high degree of accuracy compared to other whole genome sequencing methods (Liu et al., 2012). In doing so, we sought to employ SNP calling on the resulting transcriptomic data in order to identify new candidate A-to-I RNA editing sites within the rat transcriptome.

In addition to identifying new sites, we wanted to investigate known RNA editing targets that are conserved between rats and humans. Using cerebral cortex from rats at each time point, we aimed to identify whether there is a decline in RNA editing for any conserved editing sites in the rat. In particular, we wanted to identify whether the age-related decline in RNA editing of *Cyfip2* in humans was also conserved in the rat. Furthermore, several more conserved editing sites were investigated to account for the gene-specific changes observed in humans.

These results were later published in *Biogerontology* (Holmes et al., 2013).

2.3 Materials and methods

Rat tissues used in this study were supplied from a previous experiment (Merry et al., 2008). All animal husbandry procedures undertaken in this study were carried out in accordance with the provisions of the United Kingdom Animals (Scientific Procedures) Act 1986. Male BN rats SubstrainBN/SsNOlaHSD) were obtained from Harlan UK at 21–28 days of age and maintained under barrier conditions on a 12-h light: 12-h dark cycle (08:00–20:00). The health status of the rats was monitored at regular intervals through the screening of sentinel animals. All rats were fed *ad libitum* and sacrificed at 6, 12, and 28 months of age. None of the animals exhibited any signs of pathology when sacrificed. Each age group had six rats, from which brain samples were taken, flash frozen, and stored at −80°C.

RNA was extracted from cerebral cortex of rats using the RNeasy lipid tissue kit (Qiagen). The quality of the extracted RNA was assessed using the Agilent 2100 Bioanalyzer; all RNA integrity numbers (RINs) were above 8.0, indicating that the RNA had minimal degradation. For RNA extraction from brain tissue, RIN>8 represents a high quality threshold (Bettscheider et al., 2011). The samples were pooled in pairs (leaving 3 samples per age group). Ribosomal RNA was removed from the pooled samples using the Eukaryote Ribominus Kit (Invitrogen) and confirmed with the Agilent 2100 Bioanalyzer.

RNA-seq data was generated by SOLiD sequencing (Applied Biosystems) from these samples in a previous study (Wood et al., 2013). The RNA-seq results from the SOLiD system were output as colorspace FASTA and quality files. These were converted into FASTQ format using a Python script from Galaxy (http://main.g2.bx.psu.edu/). The FASTQ files were mapped to the Ensembl release 65 rat reference genome (RGSC 3.4 assembly, May 2010 gene build)

using Bowtie (Langmead et al., 2009) and settings appropriate to SOLiD data. For each sample, approximately 33.6 million reads were generated (range, 29.5 to 39.8 million reads). On average 16.7 million reads per sample were mapped to the reference genome (range, 13.8 to 21.4 million reads, approximately 50% of reads generated were mapped). All data have been submitted to GEO under the accession GSE34272 (Wood et al. 2013). A mismatch analysis was performed on the aligned reads using Bambino in order to generate candidate editing sites (Edmonson et al., 2011). These candidates were selected using custom Python scripts for A-G mismatches within exons in genes on the positive strand, and T-C mismatches on the negative strand. Results were narrowed by selecting for non-synonymous mismatches and for those with a high number of read counts. mRNA was reverse transcribed using M-MLV reverse transcriptase (Invitrogen) and amplified by PCR for selected candidate editing sites using rTaq (TaKaRa). Editing levels were then analysed by Sanger sequencing using the 3730 DNA Analyzer (Applied Biosystems) and quantifications were calculated by peak height analysis (Eggington et al., 2011).

2.4 Results

Cerebral cortex from 6, 12 and 28 month old rats were studied by RNA-seq to identify transcriptomic differences that occur during aging (Wood et al., 2013). These data were used to identify mismatches between the sequenced reads and the genome. The cerebral cortex forms the outer layer of the cerebrum and has roles in memory, language, consciousness and thought processes. Furthermore, it has been used as a platform to study RNA editing in a previous large scale study and is therefore an appropriate region to analyse (Blow et al., 2004). The reads generated were mapped to the 4.61 Ensembl build of the rat genome using Bowtie and TopHat (Langmead et al., 2009; Trapnell et al., 2009). Due to the nature of A-to-I editing, it is possible to identify RNA editing targets by isolating reads with A/G mismatches between the genomic sequence and the cDNA read sequences (or T/C for negative stand hits). Using Bambino (Edmonson et al., 2011), we identified mismatches between the transcriptome and genome, which could potentially be RNA editing targets. Using self-created Python scripts, A/G positive stand mismatches and T/C negative strand mismatches were isolated. Due to RNA-seq largely generating sequences from exons, it was decided to focus on exonic editing sites. Candidate RNA editing targets were chosen based on a number of factors, including a high read count for the site, a high percentage of the G or C (potentially edited) base at the site, and a lack of other nearby mismatch types that may indicate

sequencing or mapping errors. Sanger sequencing was then employed to analyse cerebral cortex cDNA to identify whether evidence of RNA editing mismatches would be identified by this method. Eleven genes were tested for Sanger sequencing analysis, plus a positive control, *Kcna1*, which has a known conserved editing site. This analysis led to the identification of a new editing site in rats in the Conserved oligomeric Golgi complex subunit 3 (*Cog3*) at position Chr15:61477456 (rn5 rat genome build). This modification results in a codon change from AUU to IUU, and an amino acid change from isoleucine to valine. Editing of this site has previously been reported in humans (Shah et al., 2009), and more recently in mice (Danecek et al., 2012), demonstrating conservation of the editing site between humans and rodents (Figure 2.1).

Figure 2.1: Conservation of the editing site within *Cog3* in humans, mice, and rats The genomic sequence for *Cog3,* which is expressed from the negative strand. The edited site is highlighted in red.

Using Sanger sequencing, editing levels of known conserved targets of A-to-I RNA editing were analysed from the cerebral cortex of rats used in the RNA-seq: *Gabra3, Cyfip2, Kcna1, Flna,* and *Blcap* (which has three editing sites, known as Y/C, Q/R and K/R), and the newly identified site in *Cog3*, using primers given in Table 2.2. The RNA editing enzymes, ADAR and ADARB1, have different specificities for base sequence and structure of their target sites (Riedmann et al., 2008). Therefore, different targets of RNA editing may be edited by both enzymes, or preferentially by just one. These differences have previously been hypothesized as a possible cause of the gene-specific RNA editing changes observed during aging in humans (Nicholas et al., 2010). The RNA editing sites studied here are targeted to different degrees by ADAR and ADARB1. For example, while *Cyfip2* and *Flna* are targeted primarily by ADARB1 (Riedmann et al., 2008), *Gabra3* is targeted by both ADARs (Ohlson et al., 2007), and *Blcap* is targeted mostly by ADAR (Riedmann et al., 2008). Therefore, it should be possible to identify whether any protein-specific age-related changes in RNA editing occurs in the RNA editing targets that were studied.

Gene name	**Oligonucleotide sequences**
Gabra3	*Fwd* 5'-TGTCACAAGTTTCTTTCTGGCTTA-3'
	Rev 5'-TACCTTCTTGCCTTCCCAAG-3'
Cyfip2	*Fwd* 5'-ATGGGCTTTGGCCTCTATCT-3'
	Rev 5'-ATGTTGTACTGGGGGCTGAT-3'
Kcna1	*Fwd* 5'-ATGAGGGAGTTAGGGCTGCT-3'
	Rev 5'-GATCAGTTGCGGTGCAGTTA-3'
Flna	*Fwd* 5'-AAAGGATGGCTCTTGTGGTG-3'
	Rev 5'-CTATGCACCTTGGCATCAAT-3'
Cog3	*Fwd* 5'-CTTCACGGGATGTTGTATCC-3'
	Rev 5'-TGAACTCCTCCAGTGGCTCT-3'
Blcap	*Fwd* 5'-AGCTCCTGGAGAGAGAGTCG-3'
	Rev 5'-AGCAAGTAGAAGCCCATGAA-3'

Table 2.2: Oligonucleotide sequences used for PCR amplification Oligonucleotides used for sequencing of the RNA editing targets, *Gabra3*, *Cyfip2*, *Kcna1*, *Flna*, *Cog3*, and *Blcap*.

The mean lifespan of the rat strain used for this study was 28.06±0.72 months (n=102), determined from a previous study (Merry et al., 2008). Therefore the ages of the rats used in this study, aged 6, 12 and 28 months represent 21.4%, 42.7% and 99.8% of the mean lifespan, respectively. We compared the editing levels of six human cerebral cortex samples from 76-80 year old individuals (mean 77.7 years) from a previous study (Nicholas et al., 2010) to the levels of editing present in 28 month old rats. The 76-80 year old human samples analysed represented 99.5% of the mean human lifespan, assuming a human life expectancy of 78.0 years (United States Census Bureau 2012). These samples were compared based on their similar sample size, ages, and percentage of mean lifespan. RNA editing levels of *Gabra3* remain similar in 28-month-old rats (93.2%±1.5) and 76-80 year old humans (91.5%±2.9). Editing levels of *Cyfip2* also remain high in 28-month-old rats (88.0%±1.3). However, editing of *Cyfip2* declines to 70.9%±5.7 in 76-80 year old humans. Thus, age-matched 28-month-old rats maintain high levels of *Cyfip2* editing, compared to 76-80 year old humans.

Although we observe differences in RNA editing between humans and rats, we do not identify significant changes in RNA editing levels of *Gabra3, Cyfip2, Kcna1, Flna, Blcap* or *Cog3* with age in rats by one-way ANOVA (Figure 2.2) or collectively by two-way ANOVA, $F(2,128)=0.430$, $p=0.652$. Importantly, no change was identified for *Cyfip2* ($r^2 = 0.021$), in contrast to humans (Nicholas et al., 2010). Therefore, RNA editing of these targets in the brown rat is maintained throughout the average adult lifespan.

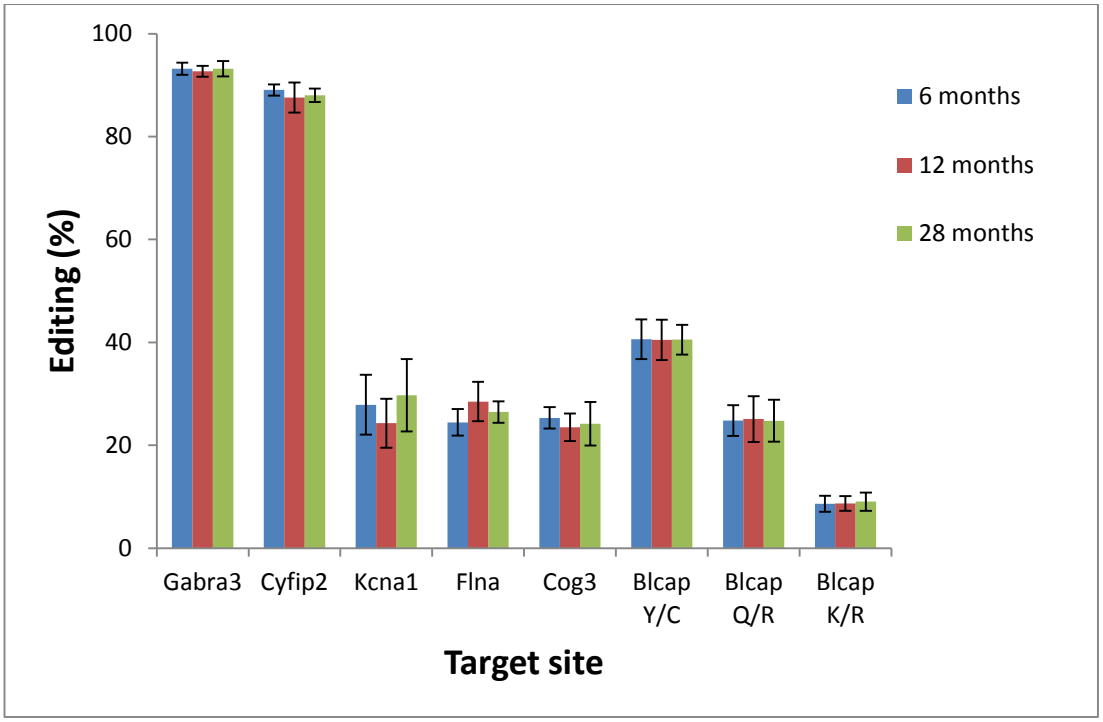

Figure 2.2: RNA editing percentages of the conserved target sites Editing percentages of *Gabra3, Cyfip2, Kcna1, Flna, Cog3,* and the Y/C, Q/R and K/R sites of *Blcap,* were quantified from cerebral cortex of rats aged 6, 12 and 28 months. Quantification was determined by Sanger sequencing and peak height analysis. Error bars represent one standard deviation from the mean. Full data given in Appendix Tables 1-4.

2.5 Discussion

We report a new editing site in the brown rat in *Cog3* which causes an amino acid change from isoleucine to valine. This site has recently been identified in humans and mice, demonstrating its evolutionary conservation (Shah et al., 2009; Danecek et al., 2012). *Cog3* is part of a protein complex in the Golgi apparatus that localises to the membrane, and is required for correct protein glycosylation (Shestakova et al., 2006). It has not yet been identified whether editing of *Cog3* can modulate its functionality, although its conservation between rodents and humans implies that it could have a functional role.

Conserved targets of RNA editing were studied in the brown rat, in order to investigate whether an age-related decline of RNA editing is conserved from humans through to the brown rat. Editing levels of *Gabra3, Cyfip2, Kcna1, Flna, Blcap* and *Cog3* were analysed in 6-, 12- and 28-month-old rats. We demonstrate that there are no significant differences in editing levels between RNA editing targets at any age in rat cerebral cortex. Unlike in elderly humans (Nicholas et al., 2010), 28-month-old rats show no deterioration of RNA editing in the targets tested in this study. Thus, within the average lifespan of the rat, we see no evidence for changes in RNA editing with age. We cannot, however, rule out the possibility that RNA editing levels may decrease in rats after the age of 28 months. In this study, only male rats were analysed, so sex specific effects cannot be ruled out. However, no differences in RNA editing have been reported to date between sexes (Zhu et al., 2012), and no differences were observed in humans (Nicholas et al., 2010).

At younger ages, humans and rats appear to show no differences in RNA editing of *Cyfip2* and *Gabra3*. In elderly humans it was identified that editing of *Cyfip2* was reduced showing a significant decline with age (Nicholas et al., 2010). In contrast, we identify that editing levels of *Cyfip2* are maintained in the rat across the mean lifespan, from 6- to 28-months of age. This suggests that the gene-specific link between RNA editing of *Cyfip2* and aging is not conserved between humans and rats. Due to the increased lifespan within the primate lineage, and since RNA editing is known to be highly prevalent in the primate lineage (Eisenberg et al., 2005b), our hypothesis is that this gene-specific link to aging may be confined to primates.

Alternatively, the decrease in RNA editing in *Cyfip2* found in humans could be due to disease pathology. Many human patients used in the previous study died from heart failure (Nicholas et al., 2010). In contrast, all the rats used in our study were sacrificed when healthy. Although there has been no report of an association between changes in RNA editing and heart failure,

it remains a possibility that this could have affected RNA editing, whether directly or indirectly.

Rodent models are frequently used in biogerontological studies, and although many aspects of aging are shared between rodents and humans, there are many important differences in their physiology, susceptibility to diseases, and their population demographics (Demetrius, 2005). Furthermore, it has recently been shown that gene expression changes in response to inflammation are considerably different between mice and humans (Seok et al., 2013). This demonstrates that not only at the physiological level, but also at the underlying genetic level, the responses of rodent models to biological processes can be very different to those in humans. Since rodents are often used as model organisms for studying aging, it is vital to understand how biological processes, such as RNA editing, are differentially affected during aging, depending on the species.

This study demonstrates that the differences in RNA editing between rodents and humans are an important consideration when using rodent models to study aging. In contrary to previous findings in humans, we report that A-to-I RNA editing of evolutionarily conserved targets does not decline with age in rats. Furthermore, we identify that the gene-specific link of RNA editing to aging, which was identified in humans, is not conserved through to the brown rat. We interpret that age-related changes observed in RNA editing could be specific to primates or humans, or that they may be a result of pathology, as opposed to aging *per se*.

Chapter 3: qPCR analysis of differentially expressed genes in the Naked mole-rat

3.1 Introduction

The naked mole-rat (NMR), *Heterocephalus glaber*, is a subterranean rodent native to Kenya, Somalia, and Ethiopia in East Africa. Its name was adopted due to its hairless appearance, although it is neither a mole, nor a rat. NMRs are one of only a few eusocial animals that live in colonies with a queen.

In living underground in colonies, NMRs are nearly blind since they have little use for eyesight in the dark. Being in an underground burrow also means that the cavern contains low oxygen levels and high carbon dioxide levels (Larson and Park, 2009), and that NMRs have adapted to these conditions. Furthermore, NMRs are less able to regulate their body temperature, with NMRs typically having a body temperature of 30-32°C (Johansen et al., 1976).

Maximum lifespan shows a positive correlation to species body size. Generally speaking, mammals with a greater body weight tend to have a greater maximum lifespan (de Magalhães et al., 2007). However, some species have a greater than expected maximum lifespan for their body size – including humans. While the predicted maximum longevity for humans, based on body size is 23.6 years, the maximum observed longevity is 122 years (Buffenstein, 2005).

At over 30 years, the maximum recorded lifespan of the NMR is the highest of any rodent species. Like humans, the NMR has a greater than expected maximum longevity for its body size (35g). Based on body size alone, NMR maximum longevity would be predicted to be 5.7 years. This is in contrast to its evolutionary relatives, mice and rats, which share a most recent common ancestor 73 million years ago (Kim et al., 2011). While mice weigh around 30g and live for a maximum of 4 years, rats weigh around 350g and live up to 5 years (Buffenstein, 2005). The NMR therefore displays a strikingly long lifespan when compared to related species.

Despite living for over 30 years, NMRs display negligible senescence, no age-related increase in mortality, and no decline in fertility (Buffenstein, 2008). They also show a high resistance to spontaneous and induced cancers (Seluanov et al., 2009; Liang et al., 2010). Resistance to cancer appears to be due to the high secretion of high molecular mass hyaluronic acid (HA)

by NMR fibroblasts. HA is a glycosaminoglycan that is a major component of the extracellular matrix. Knockdown of *Has2* which produces HA, and overexpression of *Hyal2* - an enzyme which degrades HA, both result in the loss of HA and in increased cancer susceptibility in mouse xenografts of NMR cells (Tian et al., 2013).

Therefore, NMRs display a very different ageing profile to other mammalian species – avoiding both cancer and ageing. This has made the NMR an important target of ageing research. In understanding how the NMR is able to protect itself against both cancers, and ageing processes, it may be possible to understand how these could be applied to counteracting age-related diseases in humans.

3.2 Aims

Because of the development of second generation sequencing (SGS) technologies, there has been an explosion in the amount of genetic information available to researchers, and at an increasingly low cost. In 2011, the first full sequencing of the NMR was published, identifying molecular changes in genes involved in their physiology such as thermoregulation and cancer resistance (Kim et al., 2011).

RNA sequencing (RNAseq) was employed by colleagues to identify changes in gene expression between mice and NMR. This involved mapping Illumina transcriptome sequencing reads from NMR liver to transcriptome contigs. In parallel, mouse liver was sequenced and mapped to the mouse genome. After mapping NMR contigs by BLAST to the mouse genome, it was possible to match NMR genes and their reads to mouse orthologs. From this, it was possible to identify genes that were comparatively overexpressed in the NMR when compared to the mouse.

Illumina data analysis involved mapping of reads to an incomplete NMR transcriptome consisting of contigs generated through 454 sequencing of cDNA. Mapping of the data by BLAST to the mouse genome raises the potential issue of incorrect mapping or incomplete contigs. Therefore, the aim of this analysis was to verify that genes identified as upregulated in the NMR by analysis of Illumina data were also identified as upregulated in quantitative reverse-transcription PCR (qPCR) analysis. This allowed an extra confirmation of the RNAseq results. Collectively, these findings were later published (Yu et al., 2011).

3.3 Materials and Methods

3.3.1 RNA isolation and cDNA synthesis

Tissue homogenisation of 100mg of a single naked mole-rat liver was performed using a TissueLyser (Qiagen) for 3 minutes at 25Hz. RNA was isolated by TRIzol/chloroform extraction using PureLink Mini Kit (Life Technologies) and exposed to on-column DNase I treatment according to the manufacturer's protocol. First strand cDNA synthesis was performed using 1µg of total RNA, random primers and M-MLV reverse transcriptase (Life Technologies) according to the manufacturer's protocol.

3.3.2 PCR conditions

Two microlitres of cDNA was amplified separately by PCR for the transcripts *A2m, Sat1, Sat2, Hprt1, Tbp,* and *Crym* under the following conditions: 2 minutes at 94°C, followed by 40 cycles of 30 seconds at 94°C, 30 seconds at 58°C, and 30 seconds at 72°C, followed by 2 minutes at 72°C. The following primers were used: *A2m*, 5'-GAACCGTCCTACCTCCAACA-3' and 5'-TGTTGCTGACTTCAGTTCGG-3'; *Sat1*, 5'-TGGTATAGGATCAGAAATTTTGAAGA-3' and 5'-TCCATCCCTCTTCACTGGAC-3'; *Sat2*, 5'-TCAAGGGATTGGTTCCAAAA-3' and 5'-CAGGTGGGGACAGAGATGTT-5'; *Hprt1*, 5'-GCTTCCTTCTCCGCAGACT-3' and 5'-CTTCATCACGTCTCGAGCAA-3'; *Tbp*, 5'-GAGAGGAGCTGCTTCGGATT-3' and 5'-GCTCATGCCAGAGAATAGGC-3'; *Crym*, 5'-TCCCCCAAAGACTTGAACAC-3' and 5'-CTGCCCTGAAAGAGTCTGGA-3' . PCR products were electrophoresed on a 2% agarose gel to confirm specific amplification prior to qPCR analysis.

3.3.3 qPCR

For standard curve calculations, each horizontal row of the plate was loaded with serial dilutions of cDNA after diluting the original cDNA product to 1:100. A negative control of nuclease-free water was used for each sample (Qiagen). Each sample was run in triplicate. For each primer set, a master mix was made containing 16µl forward primer (5µM), 16µl reverse primer (5µM), and 200µl Brilliant II SYBR Green QPCR Master Mix with Low Rox (Agilent Technologies). This was kept in a nuclease-free 1.5ml tube, covered in foil, mixed gently, and placed on ice until further use. 14.5µl was then aliquoted out into each well allocated to that gene. After running the qPCR, only primer sets with 93-107% efficiency (where efficiency = $-1+10^{(-1/\text{slope})}$), and $R^2>0.98$ were used for quantitative analysis.

Prior to loading the 96-well plate, a new file was created on the Applied Biosystems 7500 Fast Real-Time PCR System. Each well was labelled with the cDNA and primer set. The temperature profile was set as: 95°C for 10 seconds, followed by 40 cycles of 95°C for 30

seconds and 60°C for 60 seconds. A melt curve analysis was also performed after standard curve analysis to check for the presence of a single product.

The plate was sealed with the MicroAMP Optical Adhesive Film (Applied Biosystems), and wrapped in foil. The plate was then centrifuged at 700rpm for 2 minutes to spin the liquid down, before being placed in the machine. The cycle was then run.

Mean Ct values were calculated for those transcripts with selected primer sets. ΔCt was calculated as the difference between the mean Ct of Sat2 and the mean Ct of the selected transcript. Using the formula $N_{relative} = Ef^{\Delta Ct}$, where Ef was the mean efficiency for all six transcripts, normalised expression levels relative to Sat2 were calculated. The number of Illumina reads per transcript was normalised to that of Sat2.

3.4 Results

To verify the expression of overexpressed genes identified by Illumina RNAseq analysis, qPCR analysis was undertaken. Rather than testing the comparative levels between NMR and mice, we wanted to more directly test the NMR RNAseq results in order to confirm that the genes identified as overexpressed could be confirmed by qPCR analysis.

Primer sets were designed and tested for amplification efficiency using qPCR on serial dilutions of NMR liver cDNA to create a standard curve. Only primer sets with amplification efficiency of 93-107% and R^2>0.98 were used in this study. Six sets of primers passed this for four genes identified as overexpressed in NMR: *A2m, Crym, Sat2, Sat1,* as well two control genes, *Tbp,* and *Hprt1* (Table 3.1). An example of the standard curve generated for *A2m* is shown in Figure 3.1.

Sat2 was the most stably expressed of the genes in the qPCR data, and therefore data was calibrated to *Sat2* in order to calculate relative transcript abundances. The genes identified in the NMR liver as having the most mapped reads of the genes tested, *A2m, Crym,* and *Sat1* all showed a much greater expression than *Sat2*, which confirms the results from the RNAseq analysis (Figure 3.2). However, qPCR relative abundances show a much higher relative expression of these genes compared to *Sat2*. Indeed, expression of *Sat2* itself was the lowest of all tested genes, despite having far more reads mapping to the gene than for the control genes *Hprt1* and *Tbp* (Table 3.1). Genes identified as having a relatively low abundance from Illumina data also showed demonstrated lower abundances in the qPCR analysis (Figure 3.2). These results show confirmation that qPCR analysis is able to confirm relatively high expression for genes with a high number of mapped reads in the Illumina data analysis.

Gene	**Mapped reads**
A2m	202485
Crym	19454
Sat1	3643
Sat2	1906
Hprt1	260
Tbp	104

Table 3.1: Number of Illumina reads mapping to each gene to be tested by qPCR

The number of mapped Illumina reads does not necessarily correspond to the level of overexpression in the NMR compared to the mouse, since this depends on the number of mapped reads in the mouse RNAseq. Data from (Yu et al., 2011).

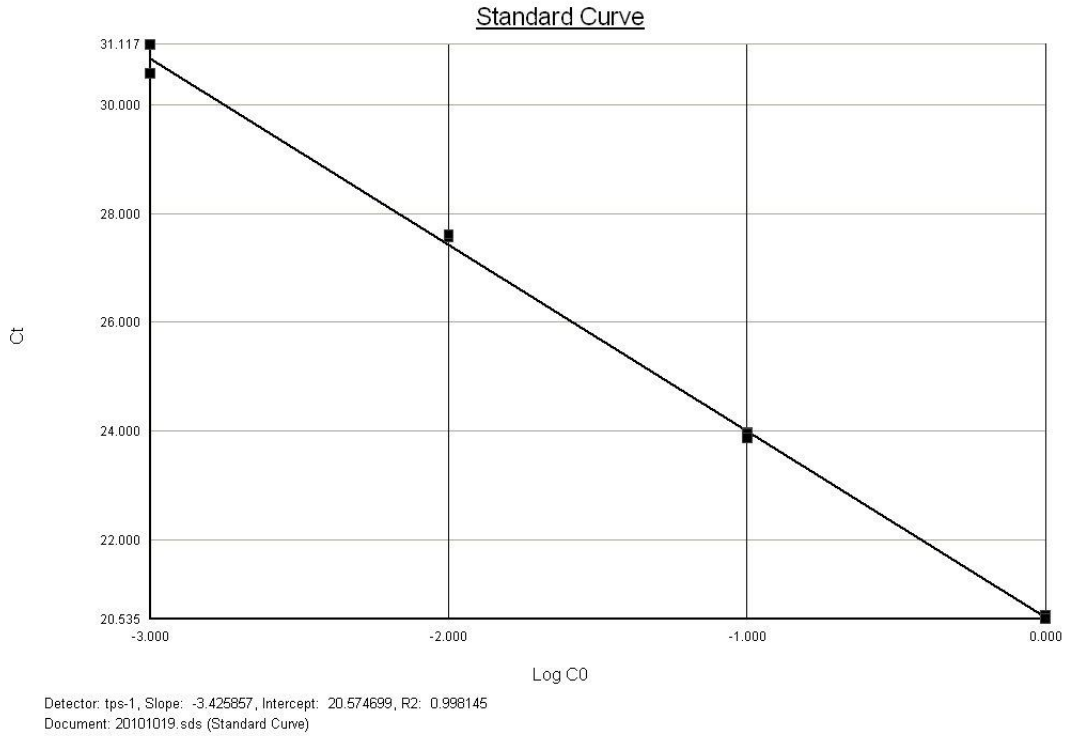

Figure 3.1: Standard curve plot for A2m

Primer sets were checked for amplification efficiency of 93-107% and a correlation coefficient of >0.98. In this example, efficiency was 95.8% and R^2=0.998. Graph generated by the Applied Biosystems 7500 Fast Real-Time PCR System.

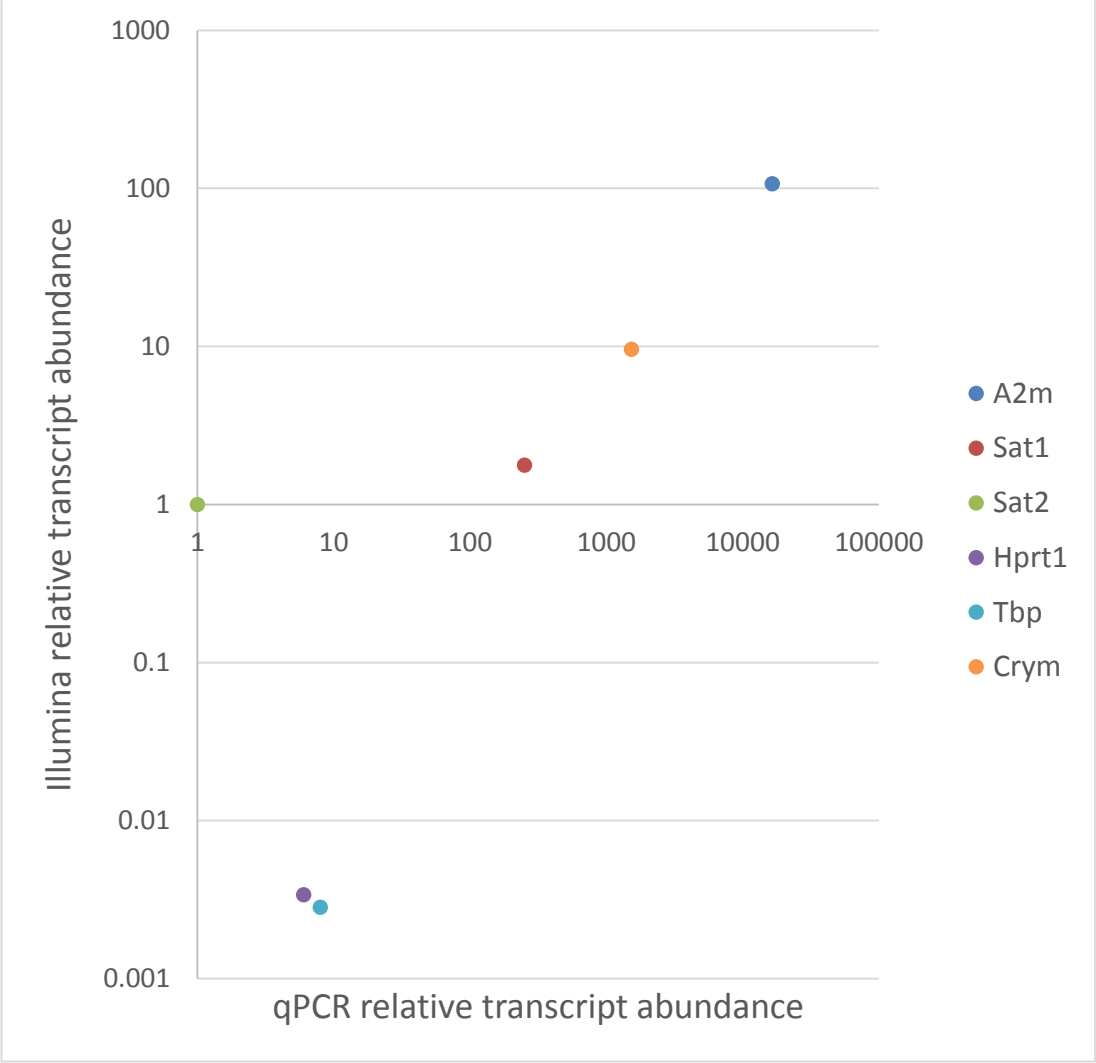

Figure 3.2: Comparison of transcript abundances between Illumina sequencing and qPCR analyses

For both Illumina mapping counts and qPCR data, relative expression was normalised to *Sat2* because of a more stable expression across repeats.

3.5 Discussion

Mice and NMRs share a most recent common ancestor 73 million years ago. Therefore, evolutionary adaptations have occurred that have led to functional changes and changes in gene family composition that make identifying orthologous genes more challenging (Kim et al., 2011).

Having an incomplete NMR transcriptome set posed a challenge in mapping of the RNAseq data. The reads were mapped to 77,086 NMR transcriptome contigs. This was a factor in identifying overexpressed genes in the NMR – since genes identified as underexpressed are likely to be genes that are not completely covered by the contigs. This would result in artificially low read counts.

After identifying genes that were potentially upregulated in the NMR, qPCR analyses confirmed that *A2m*, *Crym*, and *Sat1* were also relatively overexpressed compared to *Sat2*. Normalising to a control gene *Tbp* or *Hprt1* was not chosen due to the low read counts obtained for these genes. Because low read counts could be a consequence of poor mapping, normalising to a well expressed gene with low sample variation was deemed to be the most accurate way of comparing relative transcript abundances between RNAseq and qPCR analyses.

Global gene expression patterns of orthologous genes between mice and humans are known to be generally conserved between species (Zheng-Bradley et al., 2010). Given this, it is likely that most gene expression patterns of orthologous genes are also conserved between NMR and mice. Identifying orthologous genes that are much more highly expressed in the NMR is therefore a particularly interesting finding.

One of the most overexpressed genes identified from the RNAseq analysis was *A2m*, a protease inhibitor that has been identified as a protein that is potentially relevant to human ageing on the GenAge database (de Magalhães et al., 2009). Furthermore, *A2m* is known to interact with apolipoprotein E – which has genetic variants linked to human longevity (Garatachea et al., 2014), and is also associated with Alzheimer's disease (Blacker et al., 1998). *In vitro*, *A2m* has also been identified as a biomarker of ageing in fibroblasts (Ma et al., 2004).

Combined with previous research, this analysis identifies multiple genes that are upregulated in NMRs. These findings may provide a basis for future research on these genes in the context of NMR ageing.

Chapter 4: RNAseq analysis of *Gnasxl*$^{m+/p-}$ mice

4.1 Introduction

4.1.1 *Gnas* complex locus

One of the most documented mechanisms that can modulate ageing is dietary restriction. This involves giving animals a low calorie intake, which in many species, including mice, has been shown to significantly extend lifespan. After identifying that the decline in RNA editing with age is not universal across species, we chose to study an interesting mouse knockout of *Gnas*. These mice remain small and lean – aspects shared with calorie restricted animals. Despite this phenotype, the mice also consume more food as a result of an increased metabolic rate.

The *Gnas* gene is an extremely complex imprinted locus (Figure 4.1). With several alternative promoters, and with extensive alternative splicing, the gene encodes several transcripts, which are expressed maternally, paternally, and biallelically. Some of these are coding, and some are non-coding. Furthermore, two cis-acting antisense transcripts are also expressed from the locus, which are important for regulating the expression of *Gnas* transcripts.

Imprinting is a phenomenon whereby the expression of a gene is dependent upon which parent it was inherited from. DNA and histone methylation of the genes are different between the two alleles, which prevents expression of transcripts in a parent of origin-specific manner. Despite sequence similarities, transcripts from imprinted genes can have extremely different effects.

The major transcript encoded by *Gnas* is the G protein alpha subunit *Gαs*. This transcript is widely expressed biallelically across many tissues and cell types. However, an NH_2-terminal variant, *XLαs* is expressed solely from the paternally inherited allele, and has a more restricted expression pattern to the brain and in endocrine tissues. *XLαs* contains around an extra 300 amino acids compared to *Gαs*, containing the large *Gnasxl* exon in place of *Gnas* exon 1 (Figure 4.1).

Expression of *XLαs* is limited to neuroendocrine tissues. In the early neonatal stages of development, expression of *XLαs* is found in the SCN, PVN, DMH and VMH within the hypothalamus. In adult mice, *XLαs* is retained in these hypothalamic nuclei. Further expression is found in several distinct nuclei within the medulla of the brain, including in the solitary nucleus – a region that regulates energy control (Krechowec et al., 2012).

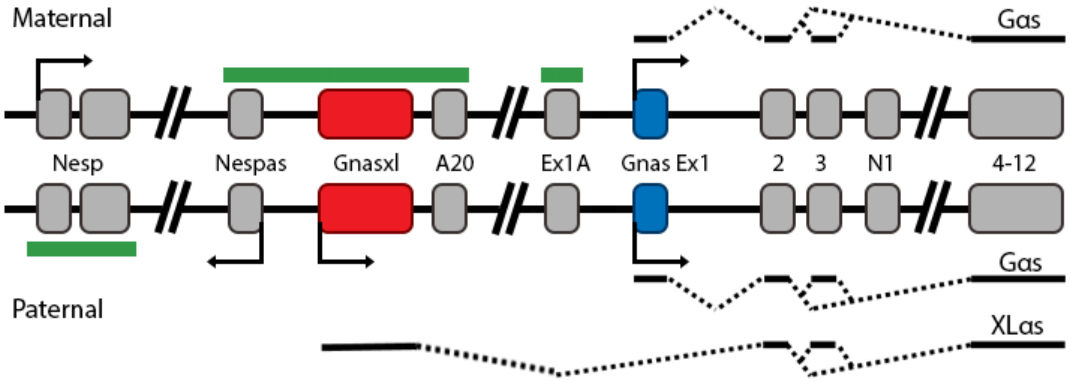

Figure 4.1: Schematic diagram of the imprinted complex *Gnas* locus

Gnas encodes two principal coding transcripts, *Gαs*, and *XLαs*. Differentially methylated regions (DMRs, shown in green) prevent transcription of exons. Although the *Gnas exon 1* promoter can be activated on both the maternal and paternal alleles, the *Gnasxl* promoter is only utilised on the paternal copy.

G proteins consist of three subunits: α, β, and γ. These interact with seven-transmembrane G protein coupled receptors (GPCRs) located in the plasma membrane to direct cell responses to a stimulus. *Gαs* is able to interact and activate adenylyl cyclase to generate cAMP for downstream signal transduction.

The G protein begins in an inactive GDP-bound state and is tethered to the membrane by the α and γ subunits (Figure 4.2A). Upon binding of an agonist to the GPCR, conformational changes result in the replacement of GDP with GTP (Figure 4.2B). The GTP-bound *Gαs* dissociates from the β and γ subunits, and activates AC, resulting in the generation of cAMP (Figure 4.2C). *Gαs* is then inactivated by GTP hydrolysis. It returns to its inactive GDP-bound state, and reassociates with the β and γ subunits (Figure 4.2D). There are several known isoforms of α, β, and γ subunits, as well as many different types of GPCRs and ACs. However exactly how these many variants and combinations affect downstream signalling is not well characterised.

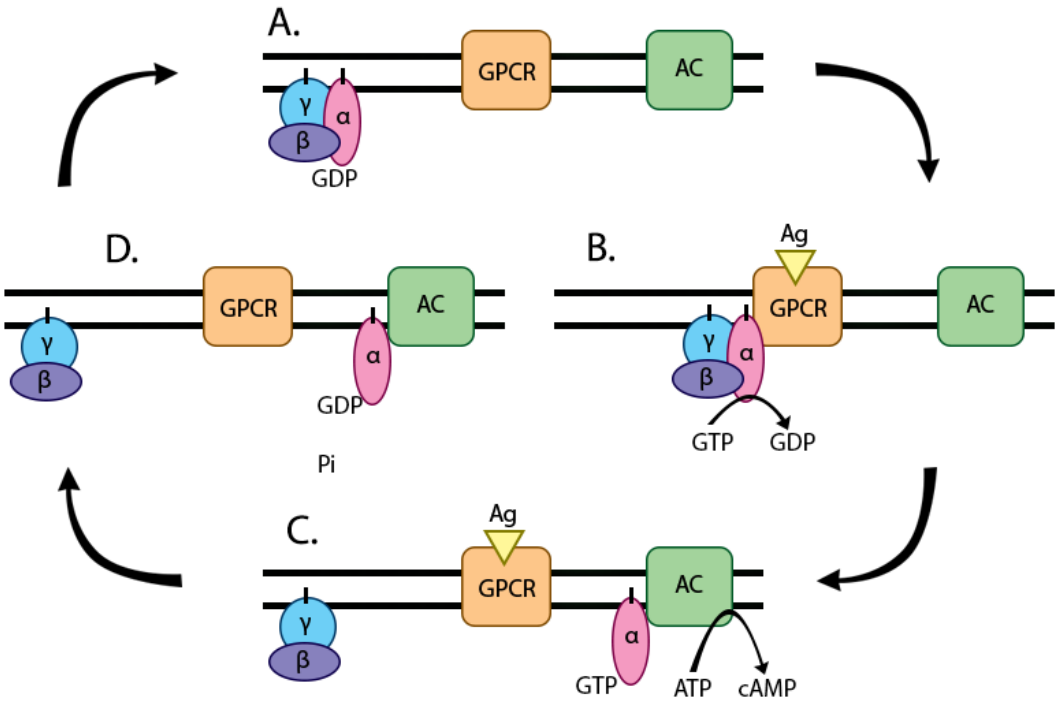

Figure 4.2: Schematic diagram of the G protein cycle

A. The heterotrimeric G protein (α, β, and γ) is localised to the membrane where it is bound to GDP. **B.** Upon binding of an agonist to the GPCR, conformational changes lead to the G protein binding GTP. **C.** The α subunit dissociates from the β and γ subunits and interacts with AC to drive the production of cAMP. **D.** GTPase-activating proteins hydrolyse the GTP back to GDP, and the α subunit reassociates with β and γ to reform the G protein.

Despite their sequence similarities, Gαs and XLαs are known to have very distinct functions. While homozygous mutations of *Gnas* are lethal, mice with heterozygous mutations can survive. Heterozygous mutations of *Gnas* exon 2 have opposing effects. When ablated from the maternal allele, mice become hypometabolic and obese (Yu et al., 2000). Conversely, when ablated from the paternal allele, mice become hypermetabolic and lean due to the loss of the paternally-expressed XLαs (Weinstein et al., 2010).

Studies have demonstrated that, like Gαs, XLαs can activate ACs to stimulate cAMP production. However, exactly how XLαs exerts different downstream effects to Gαs is not well characterised. It is known that Gαs couples with the melanocortin-4-receptor (MC4R), however it has not yet been identified which receptor XLαs couples with (Shen et al., 2013).

In vitro, XLαs has a stronger affinity than Gαs to stay localised to the plasma membrane. This is due to the presence of a highly charged amino acid domain coded by the *Gnasxl* exon. This includes two cysteine residues that can be palmitoylated to promote membrane anchoring. In doing so, XLαs is able to direct a more sustained cAMP signalling response via ACs by retaining its membrane-bound state (Liu et al., 2011).

A mouse line of *Gnasxl*-deficient mice (*Gnasxl*$^{m+/p-}$ mice) was first described in 2004 (Plagge et al., 2004). The mice were generated using a targeted deletion of a 60bp segment within the paternal copy of the *Gnasxl* exon using Cre-Lox recombinase. *Gnasxl*$^{m+/p-}$ mice have a high perinatal mortality of over 80%. Those that survive to adulthood however, remain small, lean, and hypermetabolic. *Gnasxl*$^{m+/p-}$ mice showed a significant reduction in white and brown adipose tissue mass – despite having a greater food intake. An increase in the metabolic rate was mainly due to an increase in sympathetic nervous system (SNS) activity (Xie et al., 2006). The SNS is responsible for the fight-or-flight response, but is also constantly active at a low level to regulate homeostatic mechanisms. The SNS is involved in autonomic responses such as pupil dilation, activation of sweat glands, and increasing the heart rate. More recently, it was identified that in response to the activated SNS in *Gnasxl*$^{m+/p-}$ mice, they also show a significant increase in blood pressure and heart rate, despite no significant increase in activity (Nunn et al., 2013). In addition, *Gnasxl*$^{m+/p-}$ mice showed a huge decrease in leptin levels, as well as increased sensitivity to insulin (Xie et al., 2006).

4.1.2 Diseases associated with *Gnas*

In humans, *Gnas* deficiency is responsible for several human diseases, including Albright's hereditary osteodystrophy (AHO), pseudohypoparathyroidism (PHP), pseudopseudohypoparathyroidism (pseudoPHP), and McCune-Albright syndrome. PHP is divided into three types, depending on the phenotype: Type 1A, type 1B, and type 2 (Plagge et al., 2008).

PHP is characterised by resistance to various hormones, particularly parathyroid hormone (PTH). This hormone acts to increase the Ca^{2+} concentration in the blood, whereas another hormone, calcitriol, acts to decrease it. Together, these act to retain a balanced Ca^{2+} concentration. However, due to resistance to PTH in PHP, Ca^{2+} levels decrease, leading to increased PTH and low levels of calcitriol. In addition, since PTH prevents reabsorption of phosphates in the proximal tubule of the kidney, blood phosphate levels are increased. People with type 1A PHP, also known as AHO, are also characterised by skeletal defects, such as short stature, rounded facies and brachydactyly of the fourth and fifth metacarpals, as well as a mild decrease in intelligent quotient (Garavelli et al., 2005). In addition, a lack of response to thyroid hormone leads to an increase in weight gain and an increased risk of obesity.

PseudoPHP is characterised by the same skeletal defects as Type 1A PHP, but lacks the disrupted hormone patterns, hypocalcaemia, and hyperphosphataemia. Importantly, PTH couples to the Gαs subunit to generate a cAMP-dependent response (Liu et al., 2011). Type 1A PHP and pseudoPHP are characterised by skeletal defects, obesity, and short stature, whereas type 1B PHP is phenotypically normal in this respect (Mantovani, 2011).

Although most studies on the function of XLαs have been explored in mice, evidence suggests that similar mutations in humans result in similar phenotypes. A small deletion of the *Gnasxl* exon and loss *Xlαs* expression was identified in the case of a girl displaying severe perinatal growth retardation, feeding abnormalities, facial dysmorphia, and low adipose tissue (Geneviève et al., 2005). A later study into people with PHP 1A and PseudoPHP demonstrated that patients with a paternally inherited *Gnas* mutation had a significantly low birth weight compared to those with a maternally inherited *Gnas* mutation, likely as a consequence of the loss of *Xlαs* expression (Richard et al., 2013). Therefore, studies into the function of *Xlαs* have a clinical relevance to the understanding of these diseases in humans.

4.2 Aims

Unfortunately, due to high costs for maintaining hundreds of laboratory animals over several years, it was not feasible for this project to determine whether *Gnasxl*$^{m+/p-}$ mice are long-lived. However, the mice display an interesting phenotype of reduced adipose tissue and body weight, despite increased food intake.

The hypothalamus is a critical region of the vertebrate brain that contains several nuclei that regulate many homeostatic mechanisms, such as body temperature, response to hunger and thirst, and maintaining the circadian rhythm. It is situated below the thalamus at the centre of the ventral edge of the brain. The structure is dominated by the third ventricle – a space running through the centre of the ventral side of the brain, which contains cerebrospinal fluid.

Our aims were to identify transcriptional changes in the whole hypothalamus of *Gnasxl*$^{m+/p-}$ mice that may be responsible for producing the phenotype. Due to the development of SGS technologies, there has been a rapid increase in the amount of genomic data produced, and a rapid decline in cost. Due to the advantages of rapid sequencing and generation of huge amounts of genomic data, Illumina sequencing was employed to analyse transcriptional changes within the hypothalami of wildtype and *Gnasxl*$^{m+/p-}$ mice.

4.3 Materials and methods

Hypothalamic RNA isolates were stored at -80°C. These consisted of six isolates from wildtype mice, and six from *Gnasxl*$^{m+/p-}$ mice. For quality control analysis, samples were run on the RNA 6000 Nano chip on the Bioanalyzer 2100 (Agilent Technologies). Since rRNA constitutes the vast majority of RNA, degradation of samples can be quantified by the ratio of rRNA to smaller-sized products. An RNA Integrity Number (RIN) is calculated where >8 represents high quality RNA with minimal degradation (Bettscheider et al., 2011). All of the RNA isolates showed a RIN > 8.5.

In order for the RNAseq to be more cost-effective but still retain greater biological variability, RNA samples were pooled, within conditions, in pairs. This resulted in three pooled wildtype samples and three pooled *Gnasxl*$^{m+/p-}$ samples.

rRNA accounts for around 90% of total RNA within mammalian cells (Chen and Duan, 2011). Since we are aiming to identify gene expression changes, rRNA must be removed prior to transcriptomic sequencing. Otherwise, only a small proportion of reads would map to genes other than rRNA genes. The Ribominus Eukaryote Kit for RNA-Seq (Life Technologies)

selectively pulls out rRNA using locked nucleic acid (LNA) probes. LNA is a chemically modified version of RNA, containing a methylene bridge between the 2' oxygen and 4' carbon. The sugar backbone of RNA and DNA is able to switch between two conformations. However, the methylene bridge in LNA constrains the sugar backbone into the North C3'-endo-C2'-exo conformation. Consequently, LNA has a greater affinity for RNA and allows for greater sequence specificity. These LNA probes are tagged to biotin, which has a high affinity for streptavidin-bound magnetic beads that are added to the total RNA. This allows the separation of rRNA, which is magnetically isolated using a magnetic block, while the rest of the RNA is transferred to a fresh nuclease-free tube. Since the resulting RNA is now fairly weak, it must then be concentrated using the Ribominus Concentration Module (Life Technologies).

In order to confirm the removal of rRNA from the samples, 1µl of each was run on RNA 6000 Nano chip on the Bioanalyzer 2100 (Agilent Technologies). These resulting electropherograms were compared to those prior to rRNA-depletion. From this, we confirm that the 18S and 26S rRNA peaks are much reduced, demonstrating the rRNA-depletion. A small 18S peak is retained in most samples, accounting for up to 3.2% of the RNA. This is due to the Ribominus kit being able to remove 95-98% of rRNA. Therefore some rRNA will remain even after rRNA-depletion.

The main use of the Bioanalyzer is to provide a quantitative measure of RNA quality and degradation. However since it measures the fluorescence intensity of RNA products against the migration time, it also provides estimates of sample concentration using the area underneath the curve on the electropherogram. Library preparation requires accurate quantification of RNA samples to ensure there are no biases between samples and that the cDNA synthesis and PCR reactions are achieved at optimal rates. Prior to library preparation, sample concentrations were analysed on the Qubit (Life Technologies). This uses a fluorescent dye, Ribogreen, which specifically binds to RNA, allowing a highly accurate quantification of RNA concentration. Other methods, such as the Nanodrop (Thermo Scientific) are unable to accurately quantify RNA at such low concentrations (O'Neill et al., 2011).

4.3.1 Library preparation and sequencing

For each pooled sample, 50ng of RNA was used for library preparation using the ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre). Briefly, the RNA was first fragmented at 85°C for 5 minutes, and placed on ice. cDNA synthesis primers were annealed prior to cDNA

amplification by StarScript Reverse Transcriptase at 37°C for 10 minutes. StarScript Reverse Transcriptase was inactivated by a 95°C incubation for 3 minutes. The cDNA synthesis primers consist of random hexamer sequences, with a tagging sequence attached to the 5' end. Terminal tagging of the 3' end is then performed so that the fragmented cDNA is now tagged with specific sequences at both the 3' and 5' ends. These tagging sequences allow the library reads to adhere to the sequences bound to the flow cells of the Illumina sequencing slide.

After cDNA synthesis and tagging, many impurities, such as excess primers, enzymes and salts, remain in solution with the desired tagged amplicons. The cDNA needs to be purified from these artefacts prior to sequencing to avoid interference. cDNA was purified using the Agencourt AMPure XP system (BeckmanCoulter), which uses Solid Phase Reversible Immobilisation (SPRI). These beads are paramagnetic and consist of a polystyrene core surrounded by two layers of magnetite. The surface is covered with carboxyl groups, which enables the reversible adsorption of DNA in the presence of polyethylene glycol (PEG) and sodium chloride. When placed near a magnetic microcentrifuge tube block, the DNA bound to the beads is isolated. The impurities are removed and the beads are washed twice with ethanol to ensure their removal. Finally the DNA is eluted with nuclease-free water prior to amplification.

In order to reduce representation bias during multiplex sequencing, ScriptSeq Index PCR Primers (Epicentre) were used during PCR amplification of DNA. These primers are such that the 3' end is complementary to the tagging sequence, and the 5' end contains a further tagging sequence. In the case of the forward primer, a defined six-base index sequence is placed in between the two tagging sequences. Later, during sequencing, this will allow for identification of the sample origin of the reads. The size distribution of the library preparations were checked using the Bioanalyzer 2100, and showed a profile distribution with large enough reads for sequencing (Figure 4.4). The Centre for Genomic Research (CGR) (Liverpool, UK) performed size selection on the library preparations prior to loading onto the Illumina Genome Analyzer IIx (GAIIx), which sequences reads of 100 bases in length.

The samples were loaded onto the slides by the CGR. These slides contain eight flow cell channels, referred to as lanes. Three of these were used for my samples, and five were loaded with samples from other researchers. With my samples, two were loaded per lane, and the index sequence allows for multiplexed sequencing. The lanes are coated with a lawn of primers, specific to the adaptor sequences on the library preparations of cDNA. Upon

loading of the samples, the single-stranded cDNA fragments bind randomly to the surface at both termini. Bridge amplification is performed, creating double-stranded DNA. Upon denaturation, fragments dissociate from the free terminus. These are amplified and allowed to re-bind to the adaptor sequences. Rounds of amplification generate clusters of cDNA that has been amplified from a single fragment. This is done to ensure that the fluorescent signal given off will be amplified to a detectable level such that the Illumina GAIIx machine can identify read sequences.

After sequencing was completed, the data was extracted and subjected to quality control by the CGR. This included a filter for CHASTITY ≥ 0.6, which identifies clusters with a low signal to noise ratio. This is often the result of two clusters being so close to each other that the signals cannot be distinguished from one another. Resulting data was demultiplexed using the Illumina CASAVA 1.8.2 pipeline by the CGR. Reads that passed the CHASTITY filter were filtered for, producing FASTQ files separately for the sense and anti-sense reads. A quality control analysis was then performed using FastQC (Babraham Bioinformatics) by the CGR.

4.3.2 Mapping and differential expression analysis

TopHat is a bioinformatics tool that maps transcriptomic reads to a genome. It incorporates a program, Bowtie, which performs mapping. TopHat has the advantage of calculating exon junctions based on a GTF (Gene Transfer Format) file, which contains the locations of exons, transcripts, and genes across a genome. In doing this, TopHat is able to map reads that contains exon-exon boundaries. TopHat takes the FASTQ files and maps them to the genome and transcriptome data, and outputs a file containing the sequences and the location in the genome of where they have matched.

Both sense and anti-sense FASTQ files were mapped to the mm9 build of the *M. musculus* genome (obtained from UCSC genome browser) using TopHat 1.4.1. Inner distance between mate pairs was set to 280 to account for the mean fragment size of 400 base pairs, calculated by the CGR by analysis using the Bioanalyzer 2100. Data was mapped using default settings to Cufflinks 1.3.0 to the GTF transcript annotation file provided by Illumina iGenomes. Data was later remapped using Cufflinks 2.0.2 to the GTF transcript annotation file provided by Illumina iGenomes.

After the release of the mm10 build of the *M. musculus* genome, data was reanalysed. The original FASTQ files were mapped to the mm10 genome build using TopHat 2.0.4, incorporating a repeatmask, which filters out known repeat sequences to improve mapping

quality. The repeatmask file was obtained from RepeatMasker (Tarailo-Graovac and Chen, 2009). Data was then reanalysed using Cufflinks 2.0.2 using the same settings as previously

4.4 Results

4.4.1 Quality Control and Library Preparation

In order to identify transcriptional changes within the hypothalami of wildtype and Gnasxl$^{m+/p-}$ mice, brains of six male adult mice were extracted for each background. The hypothalami were dissected out, and the RNA was isolated.

To determine the quality of the RNA, samples were run on the Bioanalyzer 2100 (Agilent Technologies) using the RNA 6000 Nano kit (Agilent Technologies). The Bioanalyzer measures RNA quality using the RNA integrity number (RIN), which is computed by comparing the amount of small products to large products, as well as the intensities of the 18S and 28S rRNA peaks. An excess of smaller products is a hallmark of RNA degradation and results in a low score. Sequencing highly degraded RNA of low quality results in poor quality sequencing, so it is imperative to ensure that RNA quality is thoroughly assessed prior to library preparation.

The resulting plot of one of the samples illustrates the distribution of RNA product sizes after isolation (Figure 4.3). The 18S and 28S rRNA peaks are also observed. The high peaks compared to the low fast region demonstrates the low levels of degradation of the RNA sample. Samples with RIN>8 are deemed to be of high quality with minimal degradation. All of the samples used in this analysis showed a RIN>8.5 (Table 4.1).

In order to maintain biological variation but keep costs manageable, samples were pooled in pairs (Table 4.1). This resulted in 3 pooled wildtype samples and 3 Gnasxl$^{m+/p-}$ samples.

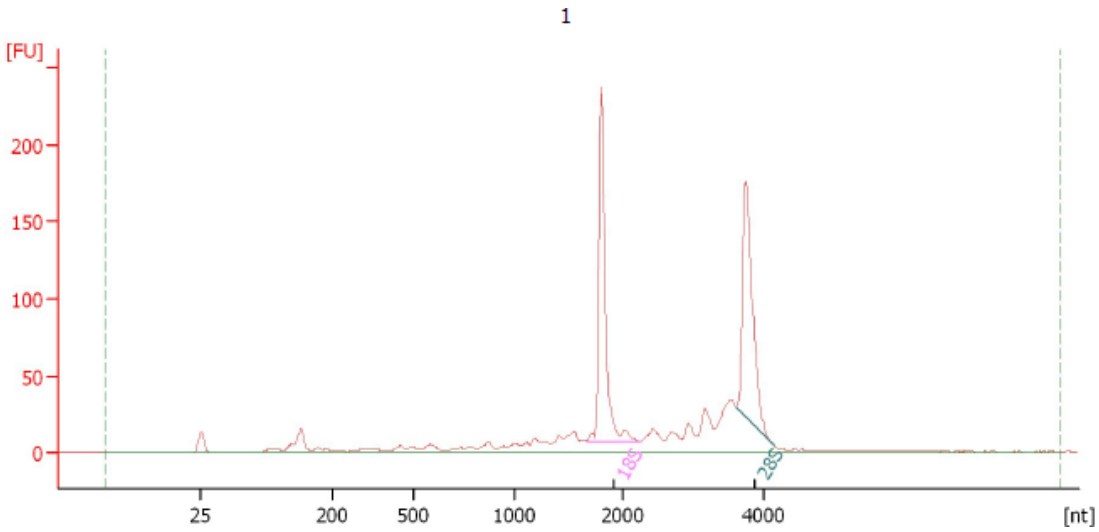

Figure 4.3: Fragment size plot by 2100 Bioanalyzer (Agilent Technologies) for sample WT-1.

The sample was run on the RNA 6000 Nano chip. Horizontal axis (nt) denotes the nucleotide length, and the vertical axis (FU) denotes the fluorescent units (measure of the intensity). 18S and 28S rRNA peaks are automatically identified by the software. The peak at 25 bases represents the marker that allows quantification.

Sample	RIN	Pooled sample
WT-1	8.6	pWT-1
WT-2	8.7	
WT-3	8.6	pWT-2
WT-6	8.5	
WT-4	8.6	pWT-3
WT-5	8.5	
KO-1	8.7	pKO-1
KO-2	8.6	
KO-3	8.9	pKO-2
KO-6	8.7	
KO-4	8.8	pKO-3
KO-5	8.7	

Table 4.1: RIN numbers calculated by the 2100 Bioanalyzer

The high RIN calculated by the 2100 Bioanalyzer indicated the quality of the isolated RNA. WT = wildtype mice; KO = Gnasxl$^{m+/p-}$ mice. Samples that were pooled together were littermates.

rRNA accounts for around 90% of total RNA within mammalian cells (Chen and Duan, 2011). Since we are aiming to identify gene expression changes, rRNA was removed prior to transcriptomic sequencing in order to enrich the mRNA content using the Ribominus Eukaryote Kit (Life Technologies). In order to confirm the removal of rRNA, samples were run on the Bioanalyzer 2100 (Agilent). These resulting electropherograms were compared to those prior to rRNA-depletion (Figure 4.4). From this, we confirm that the 18S and 28S rRNA peaks are much reduced, demonstrating the rRNA-depletion. A small 18S peak is retained in most samples, accounting for up to 3.2% of the RNA. This is due to the Ribominus kit being able to remove 95-98% of rRNA. Therefore some rRNA will remain even after rRNA-depletion. However, the RNA is now heavily enriched for mRNA.

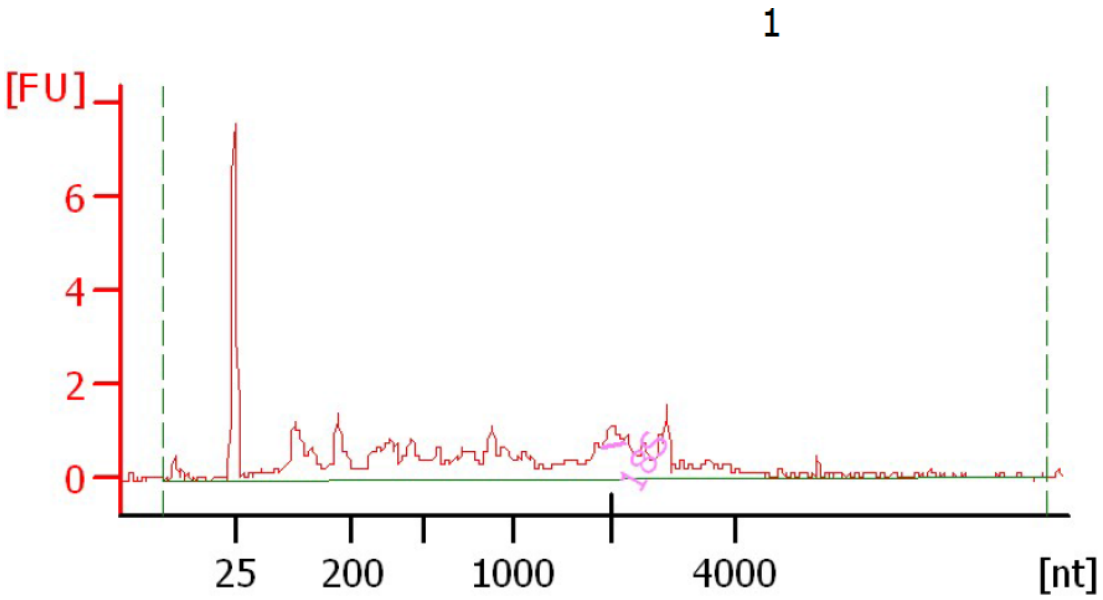

Figure 4.4: **Plot generated by the 2100 Bioanalyzer (Agilent Technologies) after the rRNA-depletion of pooled sample pWT-1**

The sample was run on the RNA 6000 Nano chip. Horizontal axis (nt) denotes the nucleotide length, and the vertical axis (FU) denotes the fluorescent units (measure of the intensity). Notice the difference in the scale of the y-axis between this and Figure 4.3).

Prior to library preparation, sample concentrations were analysed on the Qubit system (Life Technologies) in order to accurately quantify the amount of RNA to ensure equal initial RNA quantities of all samples prior to library preparation. This uses a fluorescent dye that specifically binds to RNA, allowing a more accurate quantification of RNA concentration (O'Neill et al., 2011).

Library preparation was performed using the ScriptSeq v2 Library Preparation Kit (Epicentre). The RNA was fragmented and reverse transcribed to cDNA using random hexamer primers with a known 5' tag. A terminal tag was then annealed to the 3' end, producing a cDNA fragment flanked by 3' and 5' tagging sequences. The cDNA was then purified from the solution and was amplified by PCR using ScriptSeq Index PCR Primers (Epicentre). These contain a unique tagging sequence, which is used to identify the sample after sequencing.

Samples were run once again on the Bioanalyzer 2100 to check the distribution of the size of the fragments within the library. An example of the distribution is shown in Figure 4.4 for pooled sample pWT-1. The majority of the fragments are observed to be 150-500bp in length, and therefore large enough for effective sequencing.

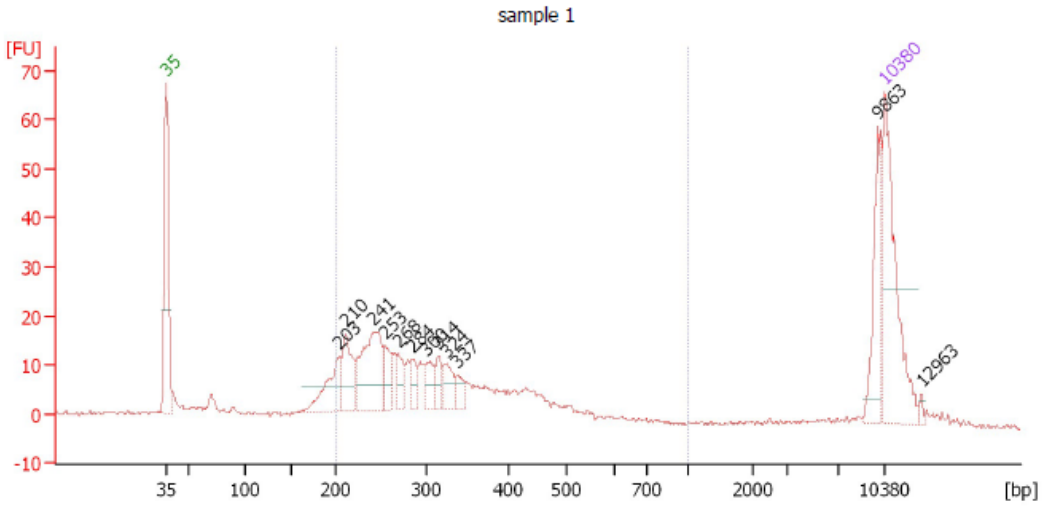

Figure 4.5: Plot generated by the 2100 Bioanalyzer (Agilent Technologies) after library preparation of sample pWT-1

Library preparation was performed using ScriptSeq v2 Library Preparation Kit (Epicentre). The sample was run on the High sensitivity DNA chip. Horizontal axis (nt) denotes the nucleotide length, and the vertical axis (FU) denotes the fluorescent units (measure of the intensity).

cDNA from the library preparation is added to the flow cell slides of the Illumina sequencing machine. The flow cells consist of channels that are coated with a lawn of adaptor sequences to which the cDNA tagging sequences can anneal to. The cDNA undergoes one round of polymerisation, and the original cDNA is removed, leaving cDNA copies attached to the flow cell. During amplification, the cDNA reads bend down and anneal to nearby adaptor sequences. This amplification creates clusters of cDNA amplified from a single read. Fluorescently-tagged nucleotides are then flowed across. After incorporation of a fluorescently-tagged nucleotide within the DNA, the fluorescent dye is released. The colour of the fluorescence for each cDNA cluster is detected by a laser and is converted to a base. This continues until all 100 bases have been read. The reads are then exported to FASTQ format, which contains the sequence data and base quality.

4.4.2 Mapping using TopHat 1.4.1 to mm9 mouse genome

The six resulting FASTQ files were mapped individually using TopHat 1.4.1 to the mm9 build of the mouse genome. Cufflinks 1.3.0 was used to calculate differential expression between wildtype and Gnasxl$^{m+/p-}$ data (Trapnell et al., 2010). Cufflinks calculates the relative transcript abundances from the aligned reads generated by TopHat. This is done by calculating the Fragments per Kilobase of exon per Million fragments mapped (FPKM). Since exons and genes have varying sizes, this needs to be taken into account to calculate relative expression values. Cufflinks is used to calculate these for each individual sample. Cuffcompare is then used to merge the results of the FPKM analysis, grouping together results from wildtype mice and those from Gnasxl$^{m+/p-}$ mice. Finally, Cuffdiff analyses the statistical differences between these groups using a multiple testing correction. Figure 4.6 outlines the analysis pipeline.

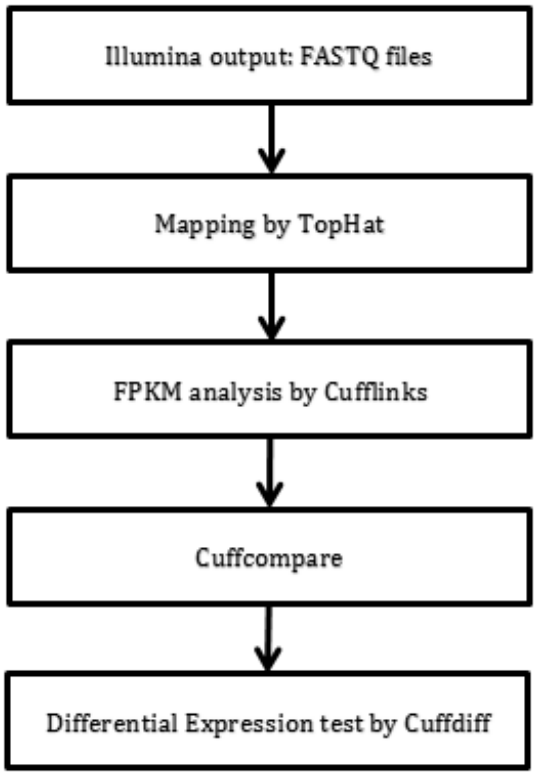

Figure 4.6: Flow diagram of the analysis pipeline after the generation of the FASTQ raw data files

Cuffdiff analysis revealed 657 genes to be significantly differentially expressed. In order to identify types of genes that were differentially expressed, the 657 gene names were input into the Functional Analysis Tool under default settings (DAVID Bioinformatic resources). Using this tool, it is possible to cluster genes into similar groups, based on similar functions, using information such as Gene Ontology (GO) terms, and protein pathways from the KEGG database. The most enriched cluster contained genes encoding proteins in the ribosome. However, many of these annotated as predicted genes. The second most enriched cluster contains genes involved in the mitochondria and oxidative phosphorylation. Eleven differentially expressed genes were identified in the NADH dehydrogenase protein complex (Table 4.2). One was identified in the succinate dehydrogenase complex. Two in the cytochrome c reductase complex, and four in the cytochrome c oxidase complex. This is an interesting finding, since the phenotype of the Gnasxl$^{m+/p-}$ mice demonstrates an overactive SNS and increased metabolic rate (Xie et al., 2006). Therefore changes to the oxidative phosphorylation protein complexes could have important effects. However the exact role that differentially expressed protein complex subunits would have on the physiology of the mice is not clear.

Since *Gnas* encodes a G-protein coupled receptor α subunit, it was also interesting to observe several G-protein related differentially expressed genes. These were identified by finding genes associated with the GO classifications GO:0007186 (G-protein coupled receptor signalling pathway), and GO:0004930 (G-protein coupled receptor activity). A total of 15 genes were identified (Table 4.3).

gene	value_1	value_2	log2(fold_change)	q_value	Ensembl gene ID	Description
Cox7c	23.843	68.9255	1.53147	2.52E-05	ENSMUSG00000017778	cytochrome c oxidase, subunit VIIc [Source:MGI Symbol;Acc:MGI:103226]
Ndufb3	9.18981	25.3052	1.46133	0.002585	ENSMUSG00000026032	NADH dehydrogenase (ubiquinone) 1 beta subcomplex 3 [Source:MGI Symbol;Acc:MGI:1913745]
Ndufa3	88.5498	236.35	1.41636	6.07E-08	ENSMUSG00000035674	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3 [Source:MGI Symbol;Acc:MGI:1913341]
Cox7a2	23.1388	54.9341	1.24739	3.02E-05	ENSMUSG00000032330	cytochrome c oxidase, subunit VIIa 2 [Source:MGI Symbol;Acc:MGI:1316715]
Ndufa1	12.3014	28.4445	1.20933	0.028788	ENSMUSG00000016427	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1 [Source:MGI Symbol;Acc:MGI:1929511]
Uqcrb	181.043	383.533	1.08302	2.65E-07	ENSMUSG00000021520	ubiquinol-cytochrome c reductase binding protein [Source:MGI Symbol;Acc:MGI:1914780]
Ndufs6	29.5758	56.574	0.935722	0.034317	ENSMUSG00000021606	NADH dehydrogenase (ubiquinone) Fe-S protein 6 [Source:MGI Symbol;Acc:MGI:107932]
Ndufa6	43.8825	77.2082	0.815109	0.004458	ENSMUSG00000022450	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6 (B14) [Source:MGI Symbol;Acc:MGI:1914380]
Cox5b	217.535	354.387	0.704079	0.004458	ENSMUSG00000061518	cytochrome c oxidase, subunit Vb [Source:MGI Symbol;Acc:MGI:88475]

Cox6c	363.943	578.964	0.66976	0.006656	ENSMUSG00000014313	cytochrome c oxidase, subunit VIc [Source:MGI Symbol;Acc:MGI:104614]
Cox6b1	226.471	130.666	-0.79344	0.002516	ENSMUSG00000036751	cytochrome c oxidase, subunit VIb polypeptide 1 [Source:MGI Symbol;Acc:MGI:107460]
Ndufb8	75.4069	41.9487	-0.84607	0.009182	ENSMUSG00000025204	NADH dehydrogenase (ubiquinone) 1 beta subcomplex 8 [Source:MGI Symbol;Acc:MGI:1914514]
Ndufv1	31.3465	17.3544	-0.853	0.002901	ENSMUSG00000037916	NADH dehydrogenase (ubiquinone) flavoprotein 1 [Source:MGI Symbol;Acc:MGI:107851]
Uqcrh	253.017	128.384	-0.97877	1.84E-05	ENSMUSG00000063882	ubiquinol-cytochrome c reductase hinge protein [Source:MGI Symbol;Acc:MGI:1913826]
Ndufa7	38.7955	18.0841	-1.10117	0.039859	ENSMUSG00000041881	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 7 (B14.5a) [Source:MGI Symbol;Acc:MGI:1913666]
Ndufa12	114.16	52.7695	-1.11328	2.57E-05	ENSMUSG00000020022	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 12 [Source:MGI Symbol;Acc:MGI:1913664]

Table 4.2: Differentially expressed genes associated with oxidative phosphorylation

Genes identified from the early TopHat 1.q value refers to the p value after multiple testing correction.

gene	Wt_value	Ko_value	log2 (fold_change)	q_value	Ensembl gene ID	Description
Cckbr	3.08702	10.0232	1.69905	0.005814	ENSMUSG00000030898	cholecystokinin B receptor [Source:MGI Symbol;Acc:MGI:99479]
Npy5r	1.83522	5.64232	1.62034	0.049822	ENSMUSG00000044014	neuropeptide Y receptor Y5 [Source:MGI Symbol;Acc:MGI:108082]
Gpr19	11.2961	22.7914	1.01266	0.009727	ENSMUSG00000032641	G protein-coupled receptor 19 [Source:MGI Symbol;Acc:MGI:892973]
Gnb4	10.027	18.9292	0.91672	0.020184	ENSMUSG00000027669	guanine nucleotide binding protein (G protein), beta 4 [Source:MGI Symbol;Acc:MGI:104581]
Atrnl1	11.458	21.3274	0.89635	0.000431	ENSMUSG00000054843	attractin like 1 [Source:MGI Symbol;Acc:MGI:2147749]
Glrb	89.5482	153.946	0.781691	0.000403	ENSMUSG00000028020	glycine receptor, beta subunit [Source:MGI Symbol;Acc:MGI:95751]
Gpr153	11.1201	18.8224	0.759281	0.040725	ENSMUSG00000042804	G protein-coupled receptor 153 [Source:MGI Symbol;Acc:MGI:1916157]
Gnai3	22.3956	36.6886	0.712114	0.014536	ENSMUSG00000000001	guanine nucleotide binding protein (G protein), alpha inhibiting 3
Homer1	28.6383	45.8556	0.679154	0.000428	ENSMUSG00000007617	homer homolog 1 (Drosophila) [Source:MGI Symbol;Acc:MGI:1347345]

Cnr1	18.1007	27.1609	0.585485	0.039168	ENSMUSG00000044288	cannabinoid receptor 1 (brain) [Source:MGI Symbol;Acc:MGI:104615]
Kcnk2	31.6309	45.1968	0.514886	6.94E-07	ENSMUSG00000037624	potassium channel, subfamily K, member 2 [Source:MGI Symbol;Acc:MGI:109366]
Cacnb4	35.2726	50.1003	0.506271	0.001069	ENSMUSG00000017412	calcium channel, voltage-dependent, beta 4 subunit [Source:MGI Symbol;Acc:MGI:103301]
Adcyap1r1	23.1971	30.117	0.376634	0.020184	ENSMUSG00000029778	adenylate cyclase activating polypeptide 1 receptor 1 [Source:MGI Symbol;Acc:MGI:108449]
Rgs7	56.7371	39.0058	-0.5406	0.01483	ENSMUSG00000026527	regulator of G protein signaling 7 [Source:MGI Symbol;Acc:MGI:1346089]
Caly	343.135	132.086	-1.37729	0	ENSMUSG00000025468	calcyon neuron-specific vesicular protein [Source:MGI Symbol;Acc:MGI:1915816]

Table 4.3: Differentially expressed genes with G-protein-related GO terms

Genes associated with the GO classifications GO:0007186 (G-protein coupled receptor signalling pathway), and GO:0004930 (G-protein coupled receptor activity). Genes from the TopHat 1.4.1 mapping to the mm9 genome build.

Interestingly, two GPCRs were identified: *Gpr19* and *Gpr153*. *Gpr19* shows expression within the hypothalamus (O'Dowd et al., 1996) and is regulated by the cell cycle (Kastner et al., 2012). *Gpr153* is widely expressed within the CNS, including in the hypothalamus where expression is found in the Arc and PVN – two regions with important roles in the regulation of feeding and appetite (Sreedharan et al., 2011). In addition, antisense knockdown of *Gpr153* in rats showed a significant reduction in food intake (Sreedharan et al., 2011).

During the RNAseq data analysis, an updated version of Cufflinks (2.0.2) was released. In addition, a major update was made to the mouse genome and annotation sets, as the mm10 version of the *M. musculus* genome was released. It was decided that data should be analysed by the latest versions in order to have the most accurate results.

After the release of Cufflinks 2.0.2, the mapped data was reanalysed using this version, generating a smaller subset of significantly differentially expressed genes. Mapping to the mm10 genome annotation in conjunction with a repeat mask and analysis using Cufflinks 2.0.2 was also performed, which also resulted in a reduced number of significantly differentially expressed genes – 289, compared to 657 in the original analysis.

Reanalysis using Cufflinks 2.0.2 generated 289 significantly differentially expressed genes, compared to 657 in the original analysis. Similarly, the numbers of differentially expressed isoforms was found to be reduced (from 828 to 289), as well as coding sequences, splicing events, and promoter usage. A smaller subset of differentially expressed genetic elements implies that the updated version of Cufflinks was more stringent in the analysis. There was also a tendency for the genes with higher p values to be no longer significantly differentially expressed after the reanalysis.

In order to further investigate the effect of the reanalysis, gene clusters were analysed by DAVID. This confirmed that the enriched ribosome and mitochondrial gene clusters were still present. Of the mitochondrial genes identified in the original analysis, 14 out of the 18 genes were still significant (Table 4,4). Four were no longer significant (*Ndufb6, Sdhd, Ndufa13, Ndufaf2*), and two new genes were identified as significantly differentially expressed (*Ndufa6, Cox5b*). In addition, analysis by the updated version of Cufflinks usually generated lower q values (after p values have been corrected for multiple testing). Combined with the result of a lower number of significantly differentially expressed genes, this suggests that the reanalysis using Cufflinks 2.0.2 was providing a more stringent analysis.

	mm9, Cufflinks 1.3.0		mm9, Cufflinks 2.0.2		mm10, Cufflinks 2.0.2	
gene	**log2(fc)**	**q_value**	**log2(fc)**	**q_value**	**log2(fc)**	**q_value**
Cox7c	1.71498	6.29E-08	1.53147	2.52E-05	1.17154	0.00972951
Ndufb3	1.50797	0.000355425	1.46133	0.00258483	1.59208	0.00190691
Ndufb6	1.42489	0.00351278				
Ndufa1	1.38604	0.00151548	1.20933	0.0287879		
Cox7a2	1.20697	3.88E-06	1.24739	3.02E-05	1.1815	1.51E-05
Ndufa3	1.17311	1.29E-06	1.41636	6.07E-08	1.14575	0.000159179
Uqcrb	0.952088	1.03E-06	1.08302	2.65E-07	1.09706	1.34E-07
Ndufs6	0.947595	0.00200254	0.935722	0.0343166	1.00665	0.0119471
Ndufa6			0.815109	0.00445776	0.695335	0.0101563
Cox5b			0.704079	0.00445776	0.657952	0.00212883
Cox6c	0.908055	3.10E-05	0.66976	0.00665577	0.65322	0.00860744
Sdhd	0.62782	0.033571				
Ndufa13	0.555144	0.0480848				
Cox6b1	-0.75371	0.00129	-0.79344	0.0025157	-0.85101	0.000142711
Ndufaf2	-0.78284	0.0347563				
Ndufv1	-0.81522	0.000789256	-0.853	0.00290132	-1.01336	0.0151469
Ndufb8	-0.92081	0.000345399	-0.84607	0.00918167		
Ndufa12	-0.96492	6.26E-05	-1.11328	2.57E-05		
Ndufa7	-1.13576	0.00922488	-1.10117	0.0398589		
Uqcrh	-1.24347	1.55E-10	-0.97877	1.84E-05	-1.06489	7.71E-07

Table 4.4: Significantly differentially expressed genes involved in oxidative phosphorylation through different analyses

Blank boxes indicate that the gene was not significantly differentially expressed. fc = fold change.

4.4.3 Mapping using TopHat 2.0.4 to mm10 mouse genome

During the reanalysis, an updated version of the mouse genome, mm10, was released. Since gene annotations were altered, data was remapped to the updated genome for the most accurate results. After mapping using TopHat 2.0.4, the data was analysed using Cufflinks 2.0.2.

The percentage of reads mapped (39.3%) to the mm10 genome was similar, but slightly higher than the reads mapped to the mm9 genome build (39.1%) (Figure 4.7). The results showed that 332 genes were identified as differentially expressed – almost half that of the original analysis (657), but slightly more genes than were identified in the second analysis (289) (Table 4.5). Despite the reduction in the number of differentially expressed genes, the various analyses showed that many differentially expressed genes were shared, as displayed by eulerAPE (Figure 4.8) (Micallef and Rodgers, 2014). Fold changes identified were mostly between a 4-fold increase and a 4-fold decrease. However, many predicted and short nucleolar RNA (snoRNAs) showed particularly high fold changes (Figure 4.9).

Two of the most upregulated genes identified in Gnasxl$^{m+/p-}$ mice were hexose-6-phosphate dehydrogenase (*H6pd)* and Hydroxysteroid 11-beta dehydrogenase 1 (*Hsd11b1)* (upregulated 4.0-fold and 3.9-fold respectively). HSD11B1 modulates the activation and inactivation of the glucocorticoid, corticosterone. Inactive deoxycorticosterone is converted to the active glucocorticoid corticosterone by HSD11B1 (Bánhegyi et al., 2009). In order to produce the active corticosterone, HSD11B1 requires NADPH. H6PD drives the action of HSD11B1 by reducing NADP+ to NADPH, while converting glucose-6-phosphate (G6P) to phosphogluconate (Atanasov et al., 2004; Bánhegyi et al., 2004; Bujalska et al., 2005). Therefore, it is an interesting finding both genes are found to be upregulated together, and suggests these genes could be acting in unison in an attempt to increase corticosterone.

	mm9, Cufflinks 1.3.0	**mm9, Cufflinks 2.0.2**	**mm10, Cufflinks 2.0.2**
Genes	657	289	332
Isoforms	828	289	403
Coding sequences	607	267	319
Splicing events	193	34	72
Promoter use	118	24	68

Table 4.5: Number of differentially expressed genetic elements through the various analyses.

All significantly differentially expressed genes identified in TopHat 2.0.4 mapping to mm10 genome and analysis by Cufflinks 2.0.2 are given in Appendix Table 5.

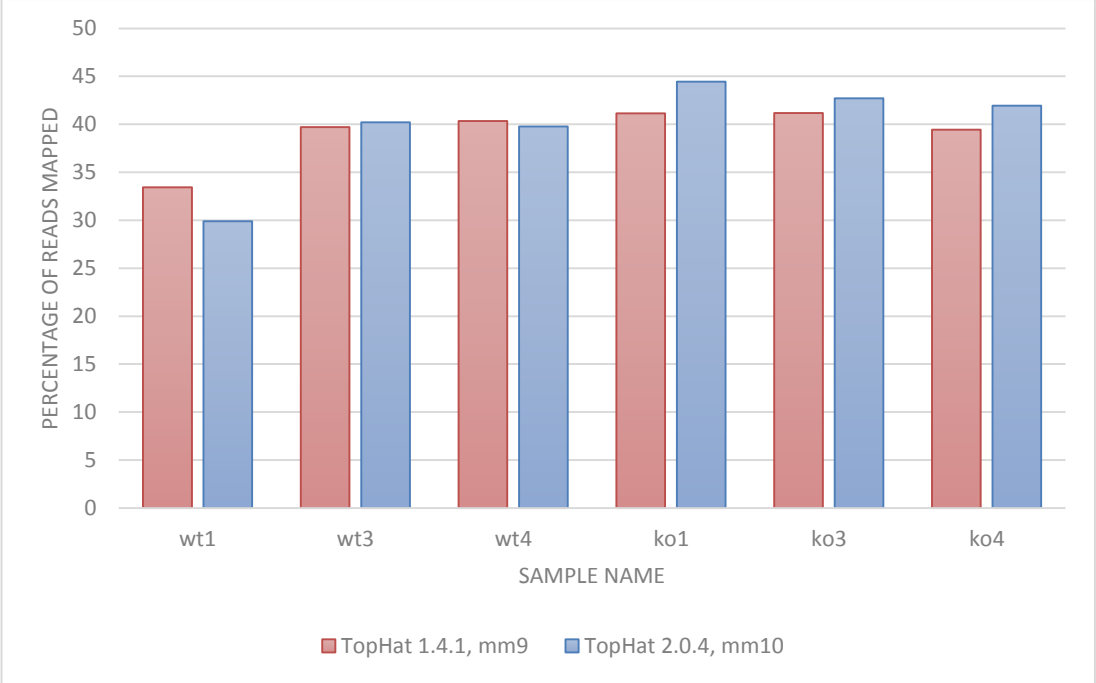

Figure 4.7: Percentage of read mapped using TopHat

Mapping percentages were similar between mapping with TopHat1.1.1 to the mm9 genome compared to TopHat 2.0.4 mapping to the mm10 genome.

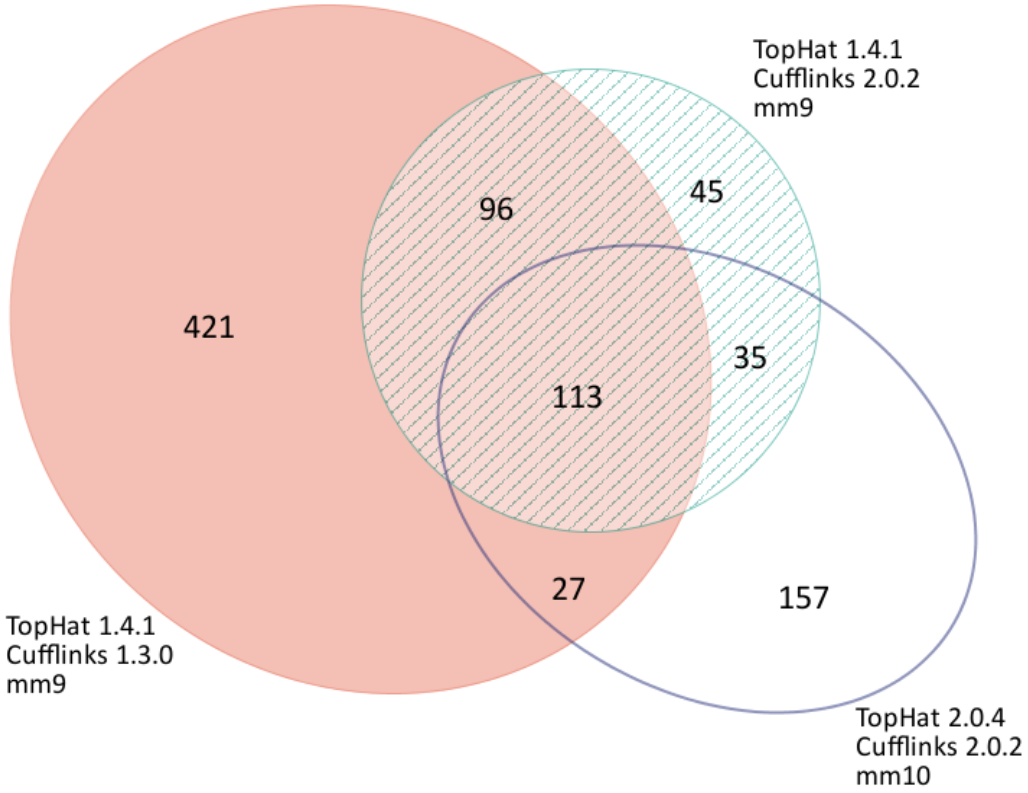

Figure 4.8: Euler diagram showing significantly differentially expressed genes that are shared between the different analyses.

A comparison between the three analyses. 422 of the 657 genes identified as differentially expressed in the original analysis (TopHat 1.4.1, Cufflinks 1.3.0, mm9) were not significantly differentially expressed based on the later analyses. Diagram was produced using eulerAPE (Micallef and Rodgers, 2014).

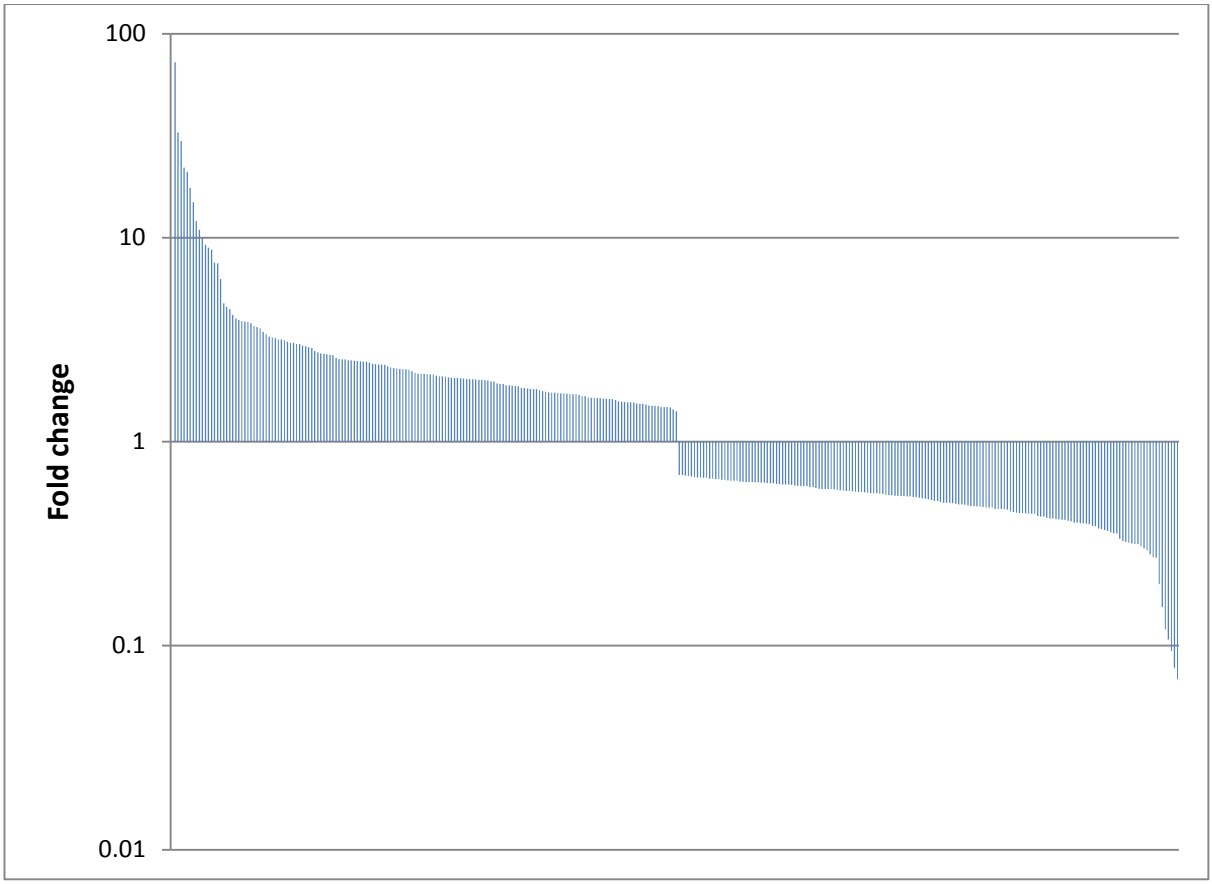

Figure 4.9: Fold changes of all significantly differentially expressed genes

Significantly differentially expressed genes from the mm10/Cufflinks 2.0.2 analysis, ordered from most overexpressed, to most underexpressed.

4.5 Discussion

These results demonstrate the importance of using the latest bioinformatic software for RNAseq analysis. Due to the increase in sequencing data, bioinformatic tools are also evolving and different versions of products can give different results. In particular, the later analyses using TopHat v2 and Cufflinks v2 suggested a more stringent analysis (Table 4.5).

Given that the bioinformatic tools are complex, there is a lot of room for discrepancies developing between different analyses. The analyses where the reads were mapped to the mm9 build of the mouse genome shared a large proportion of the significantly differentially expressed genes (Figure 4.8). However, upon mapping to the mm10 genome, almost half of the differentially expressed genes were newly identified (Figure 4.8). This may be a result of genes that were poorly annotated on the mm9 build being annotated more fully on the mm10 build. Since we did not attempt to identify *de novo* exon junctions, this will have resulted in reads spanning exons to be unmapped for poorly annotated genes. It could also be a result of using the TopHat 2.0.4 version over the TopHat 1.4.1 version and could be a result of changes in the mapping algorithms. Furthermore, the updated 2.0.2 version of Cufflinks resulted in fewer significantly differentially expressed genes than for the analysis using Cufflinks version 1.3.0.

With the increase in large scale sequencing studies, it would be interesting to investigate how analysis using different tools and inputs can affect results. Not only are there several well established mapping tools available to researchers, but there are also different quantification tools, genome builds, and sequencing platforms. Although it was not the main focus of this project, this work demonstrates the complexities involved in the analysis of RNAseq data. Bioinformatic programs are constantly evolving, but they are also limited by the input of both the quality of the data, and the quality of the genome build.

From the analysis, we identify 332 differentially expressed genes. Interestingly, some of the differentially expressed genes have been linked to roles in the regulation of energy usage, and management of body weight. Gamma-aminobutyric acid receptor-associated protein (*Gabarap*) is upregulated 2.7-fold in *Gnasxl*$^{m+/p-}$ mice. *Gabarap* is known to bind to the Angiotensin II receptor type 1 (AT1R) and increase its expression (Cook et al., 2008). Interestingly, one of the main functions of the AT1R is to modulate the SNS. Angiotensin II is able to act via the SNS to increase heart rate (Reid, 1992), which is interesting given the

recent observation that *Gnasxl*$^{m+/p-}$ mice show a significantly increased heart rate at night (Nunn et al., 2013).

Hydroxyacyl-CoA dehydrogenase (*Hadh*) was identified as being 2.7-fold downregulated in *Gnasxl*$^{m+/p-}$ mice. *Hadh* is involved in the process of fatty acid oxidation. Interestingly, knockout mice for *Hadh* show evidence of hyperinsulinism, low glucose levels (Li et al., 2010) and low body weight and decreased fat when kept on a high-fat diet (Schulz et al., 2011). Therefore, *Hadh* appears to have roles in the balance of energy regulation and in the maintenance of body weight.

Gnas was not identified as a significantly differentially expressed gene in *Gnasxl*$^{m+/p-}$ mice hypothalami. Since the *Gnasxl* exon is only expressed from the paternal copy, and having such limited expression, this was largely expected. Furthermore, no *Gnas* isoforms were identified as being differentially expressed. Because of the complexity of the *Gnas* locus and the limited expression of some exons – particularly *Gnasxl* – calculating the expression changes of rare, but biologically important, isoforms is challenging even when using the most advanced software. In addition, the transcriptomic coverage during sequencing is challenging. When expression of certain transcripts are highly localised, or expressed to very low levels, a very high coverage is necessary to detect significant changes between conditions.

Chapter 5: qPCR analysis

5.1 Introduction

Quantitative real-time reverse transcriptase PCR (qPCR) is a molecular technique that allows the relative or absolute quantification of cDNA. Essentially, cDNA is amplified by PCR and the machine detects fluorescence emitted as a product of cDNA amplification. This is done at the end of each cycle. Therefore by measuring the fluorescence, it is able to infer the amount of cDNA in the sample and assuming that the amount of cDNA product doubles each cycle, it is possible to infer the original starting amount, whether relative to another gene, or absolutely when compared to a sample of known concentration.

There are several different methods for quantifying cDNA during qPCR. One of these is SYBR Green. This is an asymmetrical cyanine dye that is able to preferentially bind within the two strands in double-stranded DNA. SYBR Green is excited under blue light (λ_{max} = 497nm) and emits green light (λ_{max} = 520nm). As it stains double-stranded DNA, the emission of green light increases as the amount of cDNA product increases during qPCR. SYBR Green is a very widely used method and remains the most cost-effective. However, its main disadvantage is that its binding to DNA is non-specific and it will bind to any double-stranded DNA product. Therefore it is extremely important to minimise secondary products. Using intron-spanning primers, which span exon-exon boundaries, prevents the amplification of any DNA contaminants in the cDNA sample. This can cause issues with simple loci, where there are few introns. Alternatively, with highly complex introns, there may not be exon-exon boundaries shared by all transcripts of the gene.

Furthermore, it is important to check for other secondary products by performing a melt curve assay. Since DNA products have different binding properties, they also have different melting and annealing temperatures. In a melt curve assay, the temperature of the sample is slowly increased, after completion of the qPCR. The qPCR machine monitors the fluorescence of SYBR Green, which decreases upon melting of DNA strands. Multiple dips will be seen in the presence of more than one product. Finally, the primer design is paramount to the qPCR. Primers must be designed in such a way that they will bind specifically to the transcript of interest, and not to elsewhere in the genome. Furthermore, they must be designed to be compatible with the amplification temperature used in the qPCR protocol; to have similar melting temperatures to each other; and to have low complementarity to each other and to themselves.

To address the potential issue of non-specific binding, Taqman probes were developed which use a probe-based method to ensure fluorescence is only emitted upon amplification of the desired product. Taqman probes consist of a fluorophore and a quencher, linked by a region that will bind a specific DNA sequence. While the probe is intact and linked to the cDNA, the quencher prevents the fluorophore from emitting light. Upon amplification, the probe is degraded by the 5' to 3' exonuclease activity of the Taq polymerase enzyme. This releases the fluorophore, which becomes excited after being separated from the quencher, and emits light. Therefore, fluorescence is directly proportional to the amount of cDNA at the beginning of the reaction. This method is more expensive than SYBR Green, and different probes have to be used for different amplicons.

5.2 Aims

Having identified 332 significantly differentially expressed genes, it was important to verify whether these results could be duplicated by reanalysis using the qPCR method. As a different method of quantification, it provides an independent analysis of the RNAseq results. Therefore, the aims of the qPCR analysis were to employ qPCR to accurately quantify a selection of genes that were identified as significantly differentially expressed between wildtype and *Gnasxl*$^{m+/p-}$ mice in the RNAseq analysis.

5.3 Materials and methods

5.3.1 cDNA synthesis

Pooled RNA samples from wildtype and *Gnasxl*$^{m+/p-}$ mice were reverse transcribed to cDNA using M-MLV Reverse Transcriptase (Life Technologies) according to the manufacturer's instructions.

5.3.2 Standard curves

For standard curve calculations, each horizontal row of the plate was loaded with five-fold serial dilutions of cDNA after diluting the original cDNA product to 1:100. A negative control of nuclease-free water was used for each sample (Qiagen). Each sample was run in triplicate. For each primer set, a master mix was made containing 16µl forward primer (5µM), 16µl reverse primer (5µM), and 200µl Brilliant II SYBR Green QPCR Master Mix with Low Rox (Agilent Technologies). This was kept in a nuclease-free 1.5ml tube, covered in foil, mixed gently, and placed on ice until further use. 14.5µl was then aliquoted out into each well allocated to that gene. After running the qPCR (see section 4.2.4), only primer sets with 93-107% efficiency (where efficiency = $-1+10^{(-1/\text{slope})}$), and $R^2>0.98$ were used for quantitative analysis.

5.3.3 Relative quantification

For relative quantification, each horizontal row of the plate was loaded with cDNA from each pooled sample. Because of using six rows instead of five, the master mix volume was increased: 20µl forward primer (5µM), 20µl reverse primer (5µM), and 250µl Brilliant II SYBR Green QPCR Master Mix with Low Rox (Agilent Technologies). The mix was treated and aliquoted identically to the standard curve protocol. Two genes, *Actb* and *Ppia* were used as control genes. Several control genes were tested, and those with the most consistent expression across samples and with the most efficient primer sets were chosen. These were later used to normalise the changes in expression across the samples to the changes in expression observed for these two genes.

5.3.4 qPCR cycling

Prior to loading the 96-well plate, a new file was created on the Applied Biosystems 7500 Fast Real-Time PCR System. Each well was labelled with the cDNA and primer set. The temperature profile was set as: 95°C for 10 seconds, followed by 40 cycles of 95°C for 30 seconds and 60°C for 60 seconds. A melt curve analysis was also performed after standard curve analysis to check for the presence of a single product.

The plate was sealed with the MicroAMP Optical Adhesive Film (Applied Biosystems), and wrapped in foil. The plate was then centrifuged at 700rpm for 2 minutes to spin the liquid down, before being placed in the machine. The cycle was then run.

Using the Applied Biosystems 7500 software, relative quantification of genes was normalised to the expression of *Actb* and *Ppia* in the pooled pWt-1 sample.

5.4 Results

SYBR Green was seen as the more appropriate choice, due its cost. In order to ensure quantification accuracy, several steps were undertaken to check for amplification efficacy and amplicon purity. To check amplification efficiency, qPCRs were run on genes against cDNA in 10-fold serial dilutions, in triplicate. C_T values were plotted against the log of relative concentration to set up a standard curve, from which the efficiency of the amplification could be tested. During a qPCR cycle, the amount of cDNA and therefore fluorescence should double each cycle, until factors such as dNTP availability become limiting.

Primer pairs were designed using the Universal ProbeLibrary Assay Design Centre (Roche), and the OligoArchitect Primer and Probe Design Solutions (Sigma-Aldrich) such that they spanned exon-exon boundaries, in order to prevent amplification of any contaminant DNA.

Primer pairs were selected on the basis that they amplified the single product at an efficient rate. This was determined firstly by the analysis of standard curve information. Efficiency is calculated on the basis of the slope of the curve, which is a measure of the amplification rate.

Efficiency must be within 93-107%, and the correlation of the points (R^2) on the standard curve must be ≥0.98 in order to ensure that the product is being amplified correctly. Secondly, a melt curve analysis was performed after the qPCR cycles. Temperature is slowly increased, and the fluorescence is measured. As the product melts, SYBR green becomes dissociated and fluorescence intensity drops. If there is a single DNA product, this will appear as a single peak. Multiple peaks indicate that there is more than one product being produced, and any primer sets amplifying these were discarded.

Several control genes were investigated, and those with the most stable expression across the six wildtype and knockout samples were chosen. *Gapdh* was considered, however it was identified as a significantly differentially expressed gene in the analysis, and therefore discarded as a control gene.

Gene	**Description**	**Efficiency (%)**	**R^2**
Ppia	Cyclophilin A	98.0	1.00
Actb	B-actin	95.6	0.99
Hadh	Hydroxyacyl-CoA-dehydrogenase	104.8	1.00
Adipor1	Adiponectin receptor 1	94.9	0.99
Gabarap	Gamma-aminobutyric acid receptor-associated protein	99.2	0.99
Rgs7	Regulator of G-protein signalling 7	103.2	0.99
Caly	Calcyon neuron-specific vesicular protein	106.4	0.98
Bex1	Brain expressed gene 1	97.9	0.99
Gfap	Glial fibrillary acidic protein	95.9	0.99
Bcl2	B-cell lymphoma 2	101.8	1.00
Mt2	Metallothionein 2	104.5	1.00

Table 5.1: Genes with primer sets used for qPCR analysis

Primer sets had an efficiency of 93-107%, and $R^2 \geq 0.98$

Gene	**Forward oligonucleotide**	**Reverse oligonucleotide**
Adipor1	CAAGCACCGGCAGACAAGA	TGGTGGGTACAACACCACTCAA
Bex1	GAGGAGAAGGCAAGGATAGGC	TCCCTTCTGATGGTATCTTGTGG
Caly	CCTCCGCATCTACACTGCAT	CCATGGTGGATGATGGTCCT
Gabarap	GGCGAGAAAATCCGAAAGAA	GATCAGAAGGCACCAGGTATTTC
Gfap	ACCAGCTTACGGCCAACAGT	CCGAGGTCCTGTGCAAAGTT
Hadh	CCTGAAGCTGAAGAACGAGCTG	TGCTGGCAAAGATGGTGTGT
Rgs7	GGCAAGGCTAGAGCTAGCAGA	GGCAAGGCTAGAGCTAGCAGA
Bcl2	TGAGTACCTGAACCGGCATCT	GCATCCCAGCCTCCGTTAT
Mt2	GCAAATGTACTTCCTGCAAGAAAAG	CAGCCCTGGGAGCACTTC
S100b	TCATTGATGTCTTCCACCAGTA	CCACTTCCTGCTCCTTGA
Actb	GCTTCTTTGCAGCTCCTTCGT	GCTTCTTTGCAGCTCCTTCGT
Ppia	CAAATGCTGGACCAAACACAA	GCCATCCAGCCATTCAGTCT

Table 5.2: Oligonucleotide sequences for primer sets

Primer sets had an efficiency of 93-107%, and $R^2 \geq 0.98$

qPCRs were run on 1:100 dilutions of the individual cDNA samples. The cDNA was amplified using RNA from the same pooled samples used in the RNAseq analysis. The samples used were prior to the rRNA-depletion. Analysis of *Adipor1*, *Bex1*, *Caly*, *Gabarap, Gfap, Hadh, and Rgs7* by qPCR suggested very little change in expression. After normalising to the expression of *Actb* and *Ppia*, the only gene to show a comparable gene expression change in the RNAseq analysis and qPCR was *Gfap*. qPCRs were repeated, but continually gave differing expression changes to the RNAseq analyses – apart from *Gfap*, where the qPCR analysis agreed strongly with the RNAseq findings.

In order to investigate the cause of the discrepancy, it was hypothesised that it could be caused by the rRNA-depletion method that was applied to the RNAseq sample, but not to the sample used for qPCR analysis. Any bias caused by the rRNA-depletion step would have huge implications for the study as it would suggest that the samples that were sequenced were not wholly representative of the total RNA that was originally obtained. Therefore, qPCRs were also run on cDNA created from rRNA-depleted RNA. Genes tested in this analysis were *Gabarap, Gfap,* and *Hadh,* as well as the two control genes *Actb* and *Ppia*. Expression changes in *Gabarap* were not significantly different to those observed in the original qPCR analysis. *Hadh* did show a slight decrease in expression in knockout samples, and more so than in the original qPCR analysis. However, the change in expression was still far from the expression change observed in the RNAseq analyses. *Gfap* on the other hand was confirmed again to have a 2-fold expression decrease in knockout mice, which was consistent with the previous qPCR analysis and RNAseq analyses.

As an intermediate filament protein that is a widely used marker of glial cells in the brain, *Gfap* has interesting characteristics. It is widely expressed in astrocytes, which could suggest changes in the number of these cells within the hypothalamus, or a decreased expression within the cells. Importantly, the presence of mRNA does not always correlate with the level of protein. In order to investigate whether GFAP was decreased at the protein level, the hypothalami of the mice were investigated in more detail.

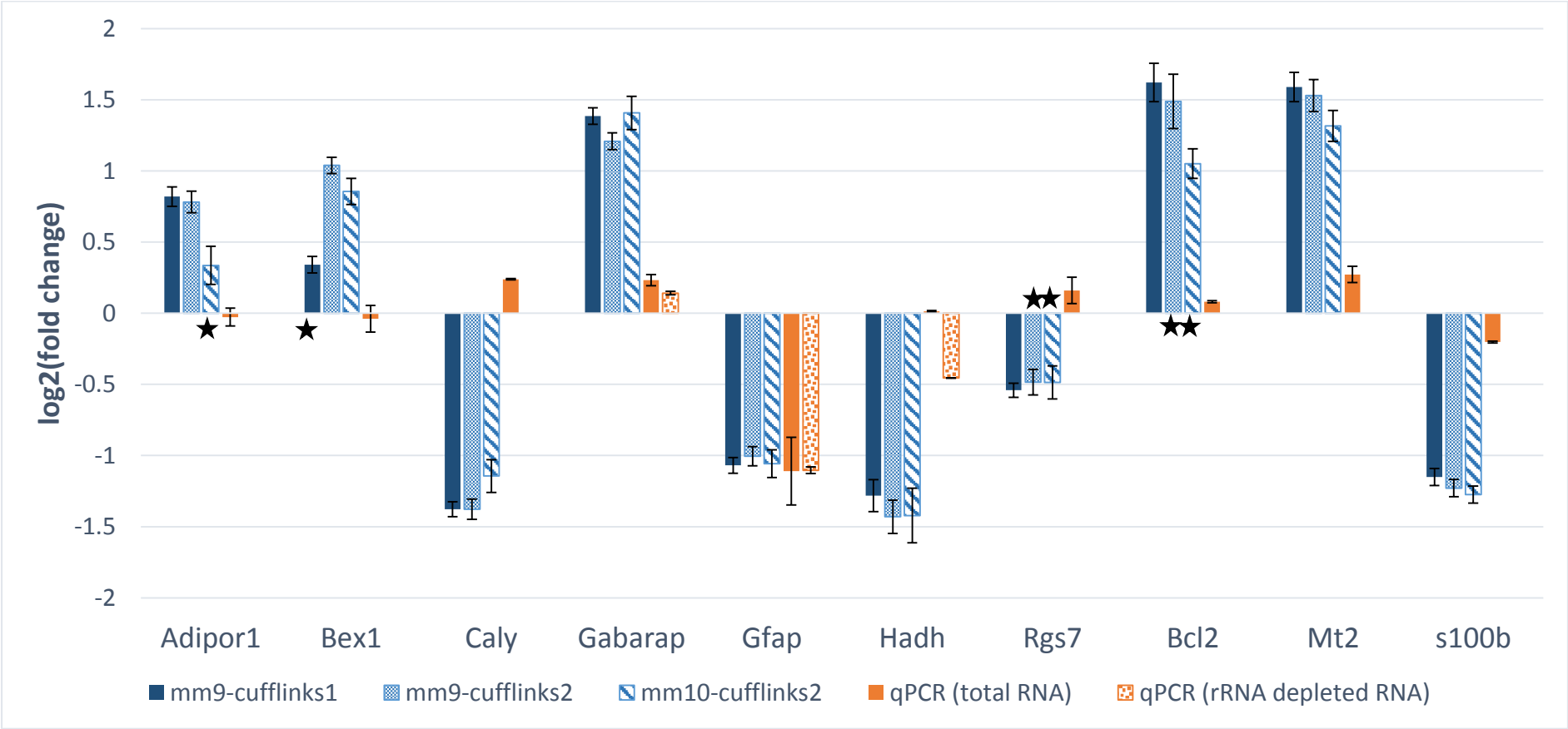

Figure 5.1: Comparison of the gene expression changes observed in Gnasxl$^{m+/p-}$ mice obtained from RNAseq and qPCR analysis

Asterisks indicate that the gene was not significantly differentially expressed in the RNAseq analysis. Error bars represent the standard error.

5.5 Discussion

After identifying 332 genes as differentially expressed between *Gnasxl*$^{m+/p-}$ and wildtype mice, we attempted to verify the authenticity of these results by performing qPCR on cDNA from the pooled total RNA samples. This allowed us to investigate whether RNAseq results were indicative of gene expression changes in the original samples.

qPCR analyses were performed on all six pooled samples with triplicate technical replicates to ensure accuracy. *Gfap* was found to be similarly decreased in expression 2-fold in both the RNAseq and qPCR analyses. However, other genes tested did not show a comparable change in expression to that recorded from the RNAseq analysis.

There can be many reasons for RNAseq analysis being discordant with qPCR results. These include PCR amplification bias during library preparation, having large standard errors – usually caused by pipetting errors, or few technical replicates in qPCRs. Because the RNA that was sequenced was from the rRNA-depleted RNA, we wanted to investigate whether there was a difference between qPCR analysis of cDNA from total RNA, and cDNA from rRNA-depleted RNA. If the rRNA-depletion step was causing a bias, perhaps by isolating some non-rRNA fragments, it could theoretically cause issues with RNAseq quantifications.

In order to investigate this, qPCRs were run on cDNA from rRNA-depleted RNA. Consequently, we identified that expression of *Gabarap* and *Gfap* was similar to the results obtained from cDNA from total RNA. For *Hadh*, rRNA-depleted qPCRs suggested a slight decrease in expression in *Gnasxl*$^{m+/p-}$ mice compared to qPCRs performed on total RNA. However, this did not match the 2.7-fold decrease identified by RNAseq (Figure 5.1). These results lead to the suggestion that rRNA-depletion did not cause undue bias in the genes that were tested by qPCR. In addition, the qPCR results generally demonstrated low standard errors for the genes that were analysed. Therefore there was very low variability between technical replicates.

It is not entirely clear why there were discrepancies between the RNAseq results and the qPCR results. It could potentially be due to biases from different amounts of starting material. However, since there appears to be a gene-specific discrepancy between the results, it could be related to the primers or the less accurate method of using SYBR Green. In either case, it would be advised for qPCR results to be quantified via another method using different primers, such as using TaqMan probes.

One of the major issues with qPCR analysis is primer design. Since SYBR Green binds to any dsDNA structure, it will also produce fluorescence when bound to secondary PCR products. Primer design is paramount in preventing non-specific amplification. In order to combat this, amplicons were run by gel electrophoresis, and by melt-curve analysis on the 7500 Fast Real-Time PCR System (Applied Biosystems), ensuring a single product of the correct size. Therefore, the primer sets passing the criteria set out in Table 5.1 only amplified the region of interest.

In response to the explosion in RNAseq experiments over the last 5 years, our work highlights the necessity of validating RNAseq results using an alternative method such as qPCR. While we identified discrepancies in expression changes between RNAseq and qPCR analyses, the 2-fold decrease in *Gfap* in *Gnasxl*$^{m+/p-}$ mice identified by RNAseq was also supported by qPCR analysis. It is not clear whether GFAP protein is decreased in the hypothalami of *Gnasxl*$^{m+/p-}$ mice. Because *Gfap* is a known marker of glial cells, the next step was to investigate how GFAP protein was affected, and whether the lack of *Gfap* suggested a change in glial cell populations, or in the amount distributed per cell.

Chapter 6: Protein analysis of GFAP

6.1 Introduction

6.1.1 GFAP

GFAP is a type III intermediate filament (IF) protein which is expressed in glial cells within the central nervous system. Proteins in this group of Ifs are able to form homo and heterodimers with one another, but not with other types of IFs. They form part of the cytoskeleton and *Gfap*, along with other IFs, are able to direct dynamic cell responses to stress and homeostasis mechanisms (Hyder et al., 2011).

GFAP expression can be induced by disease and injury to the brain. Upon brain injury, glial cells, particularly astrocytes, attempt to repair damage and fill the space left by dead cells. This is known as astrogliosis and the resulting cell mass is referred to as a glial scar. In doing so, cells proliferate and increase expression of the intermediate filament GFAP in order to alter cell morphology and extend the pseudopodia to reform the lost networked structure (Sofroniew, 2009). Interestingly, *Gfap* expression has been shown to increase during ageing and in the prevalence of Alzheimer's disease (Kohama et al., 1995; Kamphuis et al., 2014).

While GFAPα is the main isoform, several isoforms of GFAP have been identified including α, β, γ, δ, κ, and ζ (Kamphuis et al., 2012). These are of different molecular weights due to alternative splicing events. Due to the presence of different exons in each isoform, there are many antibodies that are specific to certain isoforms. The functional differences between these isoforms are not well understood. However, it has been identified that different types of astrocytes have a preference for particular isoforms. For example, GFAPδ is particularly expressed in proliferating radial glial cells within the subventricular zone, which forms the neurogenic niche (van den Berge et al., 2010). Furthermore, certain isoforms have been associated with particular pathologies such as epilepsy and Alzheimer's disease (Boer et al., 2010; Kamphuis et al., 2012).

Despite its role in responding to damage and in astrogliosis, *Gfap* knockout mice do not show an impaired astrogliosis response – possibly because of the presence of another type III IF found in astrocytes, vimentin. However, they are more sensitive to spinal cord damage (Nawashiro et al., 1998), and have an impaired response to form a protective barrier around β-amyloid, which is deposited during Alzheimer's disease (Xu et al., 1999).

6.1.2 Glial cells

Traditionally, glial cells have been thought to be merely support cells for neurons, forming the structures, but having no activity in neural transmission. However, many recent studies have identified that glia have many important roles in the CNS – they regulate the chemical environment of the brain, and they can modulate neuronal activity through synapses. In rodents, glia account for around 60% of cells in the CNS. In humans, glia account for around 90% of the cells. Recent studies have demonstrated that the engraftment of human glial progenitor cells into immuno-compromised mice increases the ability of the mice to learn (Han et al., 2013), highlighting the important roles that glia have on brain function.

Neurons have traditionally been the main focus of brain research. However, over the last decade, research into understanding the roles of glial cells has expanded rapidly (Freeman and Rowitch, 2013). In particular, there have been many recent studies implicating glial cells in the role of energy regulation within the hypothalamus.

There are three main classifications of glial cells: macroglia, and microglia. Microglia are the main immune defence mechanism in the CNS. They maintain a star-like morphology, with a small cell body, and thin pseudopodia extensions. These search for danger within the CNS, such as damaged neurons and infection. They may phagocytose viruses and bacteria and present antigens to T-cells to instigate an immune response.

Macroglia consist of ependymal glia, specialised tanycytes, oligodendrocytes, and astrocytes – the latter of which make up the largest proportion of glial cells. Most macroglia express *Gfap,* except for oligodendrocytes.

Ependymal cells line the periphery of the brain and the cerebral ventricles spaces. Specialised tanycyte cells line the wall of the third ventricle (3V) and the median eminence (ME) region within the hypothalamus. The 3V contains cerebrospinal fluid, containing nutrients that can be distributed to the parenchyma. While the cell body of the tanycytes remain in the wall of the 3V, the tanycytes send out long processes into the parenchyma of the hypothalamus. They form interactions with nuclei that are known to have important roles in the regulation of appetite, feeding, and energy regulation. Through these connections, the tanycytes are able to distribute nutrients and are hypothesised to relay signals to specific regions of the hypothalamus in order to direct responses. Interestingly, the tanycytes are able to undergo neurogenesis, producing new neurons in adult mice. Furthermore, targeted blocking of neurogenesis of the tanycytes resulted in weight gain (Lee et al., 2012). In addition, the ME region of the brain is outside of the blood-brain barrier, meaning that the tanycytes in this

region act as a bridge to relay hormonal changes to the brain. In particular, a signalling cascade is initiated in tanycytes when exposed to leptin (Balland et al., 2014). Leptin is a hormone that is released from fat-containing adipocytes. Leptin is released when the animal is full to promote satiety. Because obesity is associated with elevated leptin levels, it has been postulated that tanycytes may be of therapeutic importance for the treatment of obesity (Balland et al., 2014).

Astrocytes are one of the most common cell types within the CNS, and they are a diverse group of cell. They span the entirety of the CNS and are spaced apart in an organised manner. Typical astrocytes have a star-like morphology, and are characterised by multiple processes extending away from the cell body. They widely express GFAP and vimentin – both intermediate filament proteins. However there are some subsets of astrocytes that do not express GFAP. Astrocytes have the ability to modulate neuronal activity through synapses. There is also increasing evidence for astrocytes having an important role in the regulation of energy balance, metabolism, and appetite (Bélanger et al., 2011).

6.1.3 The hypothalamus

The hypothalamus is a region of the brain, situated at the centre ventral edge of the brain above the optic chiasm (Figure 6.1, Figure 6.2). One of its functions is to connect the endocrine system to the nervous system via the pituitary gland. Consequently, the hypothalamus is responsive to and synthesises many hormones. The hypothalamus is a critical regulator of homeostasis, in particular in regulating body temperature, the circadian rhythm, thirst, appetite, feeding behaviours, and energy expenditure via the SNS.

Several sub-regions, referred to as nuclei, comprise the hypothalamus (Figure 6.3, Figure 6.4). Each region has specific roles in responding to hormones and signals in order to manage these regulatory patterns. During development, neuronal and glial networks are established to create circuits that act to respond to feeding, appetite, and other factors governed by the hypothalamus (Bouret and Simerly, 2006). However, gliogenesis and many hypothalamic neuronal connections do not occur until early postnatal stages. The early stages of life are critical in establishing the hypothalamic networks.

6.1.3.1 Suprachiasmatic nucleus

The suprachiasmatic nuclei (SCN) are situated within the anterior part of the hypothalamus, directly above the optic chiasm (Figure 6.4). It consists of two small densely packed areas of cells that are located close to the base of the third ventricle, which runs through the centre of the hypothalamus. The primary role of the SCN is in controlling the circadian rhythm. In

doing so, the region has a very central role to play in the regulation of homeostatic mechanisms, many of which are under circadian control, such as hormone levels and body temperature regulation.

6.1.3.2 Arcuate nucleus of the hypothalamus

The arcuate nucleus (Arc) is situated at the mediobasal hypothalamus (Figure 6.3). It surrounds the base of the 3V and is characterised by its role in appetite and feeding. Connections from the Arc are minimal after birth, but begin to develop postnatally (Bouret and Simerly, 2006). By postnatal day 6 (P6), connections extend towards the dorsomedial hypothalamus (DMH). By P10, connections to the paraventricular nuclei of the hypothalamus (PVN) are established. Only by P12 do the connections extend outwards to the lateral hypothalamus (LH)

The Arc contains many types of neurons. Orexigenic neurons act to stimulate appetite, and express Agouti-related peptide (AgRP) or Neuropeptide Y (NPY). Anorexigenic neurons act to suppress appetite, and express proopiomelanocortin (POMC). They are also capable of stimulating energy expenditure via the SNS. These neurons are connected to other regions of the hypothalamus, particularly to the PVN and DMH.

Leptin is a hormone produced by adipocytes in white adipose tissue, which is produced when the adipocytes accumulate fat. Leptin circulates to the hypothalamus where it binds to leptin receptors on AgRP/NPY and POMC-expressing neurons in the Arc. On POMC neurons, leptin binding stimulates the release of α-melanocyte-stimulating hormone (α-MSH) via downstream signalling of the Jak-Stat pathway. α-MSH is able to bind to melanocortin-4-receptor (MC4R) to suppress appetite.

AgRP/NPY neurons on the other hand are repressed by leptin, but stimulated by another hormone, ghrelin. This is secreted by cells in the gastrointestinal tract when the stomach is empty. Ghrelin circulates and binds to growth hormone secretagogue receptor on AgRP/NPY neurons. This causes the release of γ-aminobutyric acid (GABA), AgRP, and NPY, which act to inhibit MC4R and therefore increase appetite.

6.1.3.3 Paraventricular nucleus

The paraventricular nuclei (PVN) consist of two clusters of cells adjacent to the top of the third ventricle (Figure 6.4). This region contains many neuronal connections of significance. Both AgRP/NPY and POMC neurons send axons from the Arc to the PVN.

6.1.3.4 Development of the hypothalamus

The progenitor cells of glia begin in the neural tube during embryonic development. Sonic hedgehog (SHH) is secreted from the notochord and the floor plate of the neural tube, and its antagonist BMP is secreted from the roof plate. This creates a morphogen gradient that allows cells to differentiate into different types of cells depending on their location and how much SHH they receive (Ingham and Placzek, 2006). This is critical in patterning the neural tube to direct the differentiation of cells into different cell types.

Cells that express *Shh* early in development, prior to embryonic day (E)9.5, go on to produce neurons, tanycytes, and astrocytes towards the posterior regions of the hypothalamus, including the Arc and ME (Alvarez-Bolado et al., 2012). Cells expressing *Shh* after E9.5 and before E12.5 give rise to neurons and astrocytes within the tuberal central region of the hypothalamus, particularly within the VMH and DMH. After E12.5, nearly all *Shh*-expressing cells contribute hypothalamic astrocytes, representing a developmental change from neurogenesis to gliogenesis (Alvarez-Bolado et al., 2012).

While neurogenesis occurs from around E8 to E20, gliogenesis at a slightly later stage, but continues postnatally for several weeks. Astrocytes begin to develop early in gliogenesis, with differentiation peaking around P2. Differentiation continues for around 2 weeks postnatally. Whereas oligodendrocytogenesis differentiate later in gliogenesis, peaking at around P14 (Sauvageot, 2002).

The subgranular zone within the hippocampus has an important role in neurogenesis. Within the neurogenic niche, radial glial cells act as a scaffold for new neurons to migrate to the granule cell layer. Radial glial cells can divide asymmetrically to produce a new stem cell, and a neuronal precursor cell. They also express *Gfap*, since they are a subset of astrocyte cells.

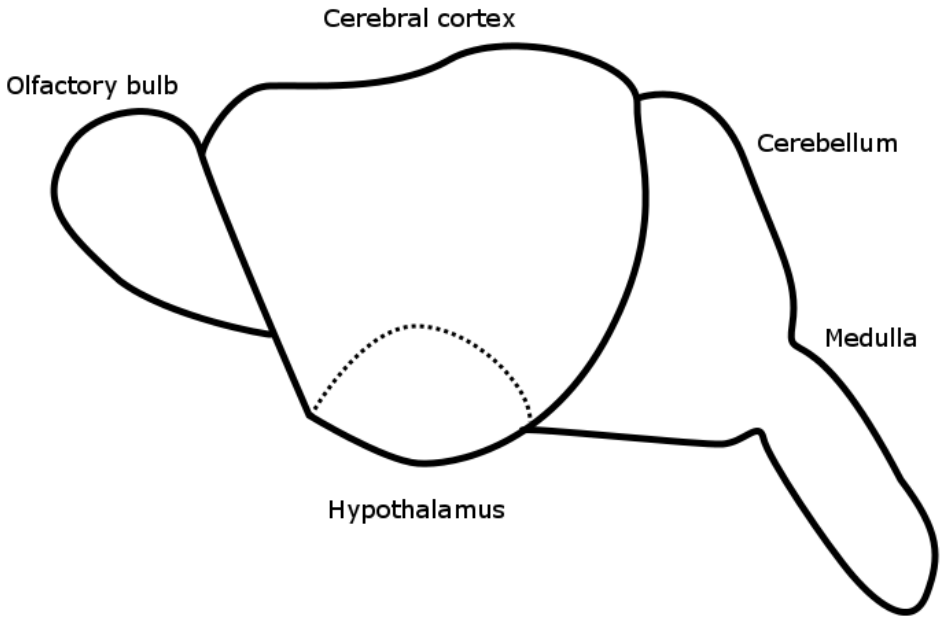

Figure 6.1: Schematic diagram of the side view of the mouse brain

The olfactory bulb is the most anterior of the brain regions. The mid-brain consists of several compartments, including the hypothalamus (indicated by dotted line), which is situated at the ventral edge of the centre of the brain. The cerebral cortex surrounds the outer portion of this region, and the hypothalamus is not visible from this angle. The most posterior regions are the cerebellum and the medulla.

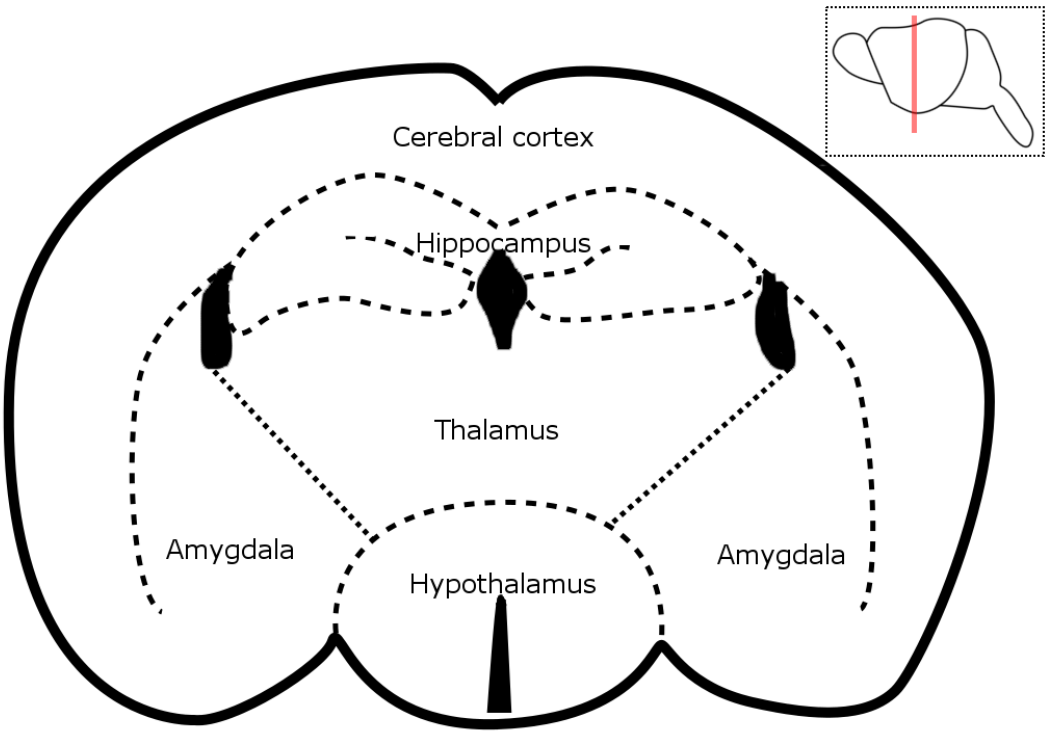

Figure 6.2: Schematic diagram of a coronal cross-section of the mouse brain

Cross-section showing the main regions at bregma -1.6mm. This is roughly half way along the length of the mouse brain. The hypothalamus is situated at the ventral edge of the brain and surrounds the third ventricle. Solid areas filled in black represent the ventricle spaces, which are filled with cerebrospinal fluid.

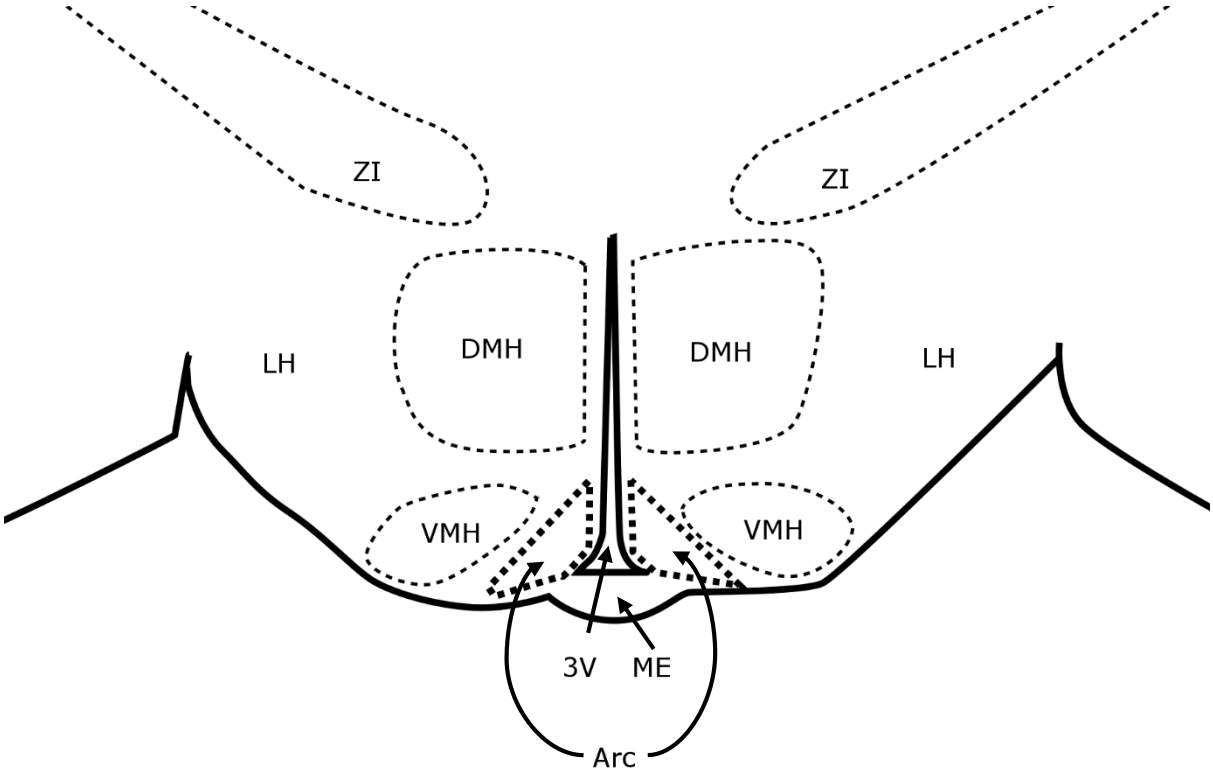

Figure 6.3: Schematic diagram of a coronal cross-section of the Arcuate nucleus

Cross section at bregma -1.94mm. ZI = Zona incerta; LH = Lateral hypothalamus; DMH = Dorsomedial hypothalamus; VMH = Ventromedial hypothalamus; Arc = Arcuate nucleus; ME = Median eminence; 3V = Third ventricle.

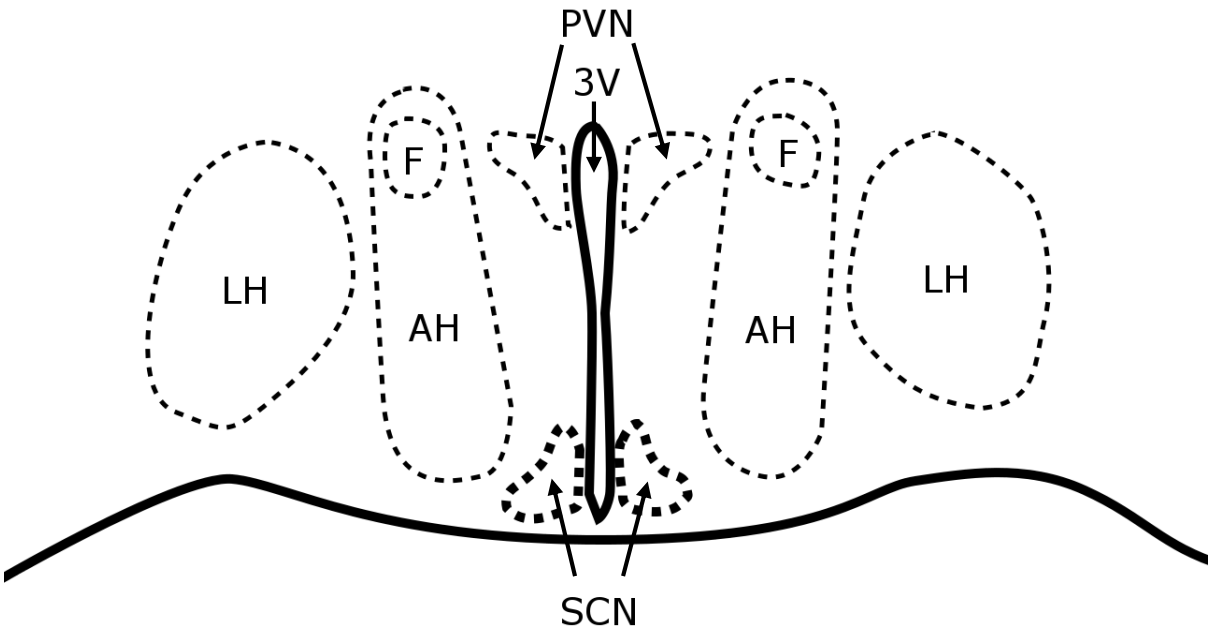

Figure 6.4: Schematic diagram of a coronal cross-section of the SCN and PVN

Cross section at bregma -0.70mm. LH = Lateral hypothalamus; AH = Anterior hypothalamus; PVN = Paraventricular nucleus; SCN = Suprachiasmatic nucleus; F = Formix, 3V = Third ventricle.

6.2 Aims

From identifying that *Gfap* was downregulated 2-fold in *Gnasxl*$^{m+/p-}$ mice in both the RNAseq and qPCR analyses, we aimed to identify whether the same was true of GFAP protein in the hypothalamus. Furthermore, because of the many specialised nuclei within the hypothalamus – many of which involved in the regulation of appetite and feeding behaviour – it was also important to investigate the localisation of GFAP. Using immunofluorescence, it is demonstrated here that *Gnasxl*$^{m+/p-}$ mice show a decline in tanycyte projections from the Arc, and fewer GFAP-positive cells in the SCN.

From these results, it was not clear whether this decline in GFAP in adult *Gnasxl*$^{m+/p-}$ mice was due to a developmental problem, or whether it was a response to the chronic undernutrition that the mice undergo as a result of an overactive SNS and hypermetabolism. In order to investigate this question, GFAP expression was also analysed in P1 and P5 mice to identify whether there is a developmental problem during gliogenesis.

6.3 Materials and Methods

6.3.1 Tissue fixation and extraction

Mice were anaesthetised by Antonius Plagge via an intraperitoneal injection of pentobarbitone. Once brain-dead, the chest was cut and the right atrium of the heart was cut in order for the blood to drain from the circulatory system. The tissues of the mice were fixed by injecting PBS into the left ventricle, to clear the system of blood, followed by 4% paraformaldehyde (PFA) in PBS. To extract the brain, the skull was carefully cut open, and the brain was dissected out. The tissues were placed in 4% PFA overnight. They were then transferred to a solution of 30% sucrose for 24-48 hours at 4°C in order to dehydrate them prior to sectioning.

6.3.2 Protein lysates

Individual hypothalami were dissected from the brain tissue and homogenised and sonicated by Katie Burton in 400µl radio-immunoprecipitation assay lysis buffer with protease inhibitor (Sigma-Aldrich), PMSF (Sigma-Aldrich) and PhoSTOP (Roche). Samples were placed on ice and centrifuged at 12000g for 10 minutes. Lysate supernatants were separated into 100µl aliquots and frozen at -80°C until further use. Using the Pierce BCA Protein Assay kit (Thermo Scientific, 23227), protein concentration was quantified according to the manufacturer's instructions.

6.3.3 Immunofluorescence

Brain tissues were carefully cut to isolate the midbrain section of interest. Sectioning of adult *Gnasxl*$^{m+/p-}$ and wildtype mice brains was performed on the Leica CM 1950 Cryostat at -20°C by Katie Burton, using Shandon Cryomatrix mounting media (Thermo Scientific). Frozen coronal sections at a thickness of 14µm were transferred to SuperFrost Plus slides (Thermo Scientific) and stored at -80°C until further use. Sectioning of postnatal day 1 and postnatal day 5 mice (P1 and P5 respectively) for both *Gnasxl*$^{m+/p-}$ and wildtype mice brains was performed by myself, using the same method.

Slides were incubated in 10mM sodium citrate for one minute at 60°C; washed three times in PBS at room temperature for 5 minutes; blocked for one hour at room temperature in 10% donkey normal serum in PBS; and incubated overnight at 4°C in 10% donkey normal serum in PBS, 0.25% triton, 1:500 Polyclonal rabbit anti-GFAP primary antibody (Dako Z0334) in PBS. Slides were then washed three times for 15 minutes in PBS at room temperature. These were incubated in 10% donkey normal serum in PBS, 0.25% triton, 1:1000 4',6'-diamidino-2-phenylindole (DAPI) and 1:1000 donkey anti-rabbit IgG Alexa488-tagged secondary antibody for one hour at room temperature in the dark. Slides were washed three times in PBS for 15 minutes prior to mounting with Fluoro-gel (Electron Microscopy Sciences), drying, and sealing with clear nail varnish. Slides were stored in the dark at 4°C.

6.3.4 Microscopy

Epifluorescence microscopy was performed using the Zeiss Axioskop 40 microscope. Images were taken and processed using the AxioVision 4.6 software (Zeiss).

6.3.5 Self-cast SDS-PAGE

6.3.5.1 Anti-GFAP antibody incubation

After transferring, the membrane was carefully removed and marked in the top corner to mark the side on which the protein adhered to. The membrane was washed in PBS prior to blocking in PBS Tween with 5% non-fat dried milk powder for one hour at room temperature. Membranes were sealed in a plastic bag and incubated with 2ml 1:6000 polyclonal rabbit anti-GFAP primary antibody (Dako Z0334) in blocking solution (PBS Tween and 5% non-fat dried milk powder) for one hour at room temperature on a shaker. Membranes were washed four times for five minutes each in PBS Tween prior to a one hour incubation in a sealed plastic bag with 10ml 1:10000 donkey anti-rabbit horseradish peroxidise (HRP)-tagged secondary antibody. Membranes were washed again in PBS Tween four times for five minutes each to remove excess antibody.

6.3.5.2 Enhanced chemiluminescence detection and exposure to film

Enhanced chemiluminescence (ECL) was performed, in order to detect presence of the HRP-tagged secondary antibody. This was achieved using the Amersham ECL Plus Western blotting reagent pack (GE Healthcare), according to the manufacturer's instructions. Briefly, excess wash buffer was drained prior to incubating the membranes in a 40:1 ratio of solution A to solution B on the protein side of the membrane, for five minutes. Membranes were then drained of excess ECL reagent and wrapped protein side down in Saran wrap. A fluorescent marker was attached to the front of the Saran wrap so that after exposure onto film, it would be possible to transfer the bands of the protein ladder to the film in order to estimate the densities of proteins within the bands present on the membranes. These were then placed protein side up in a film cassette in the dark room. The length of exposure time was varied, depending on the strength of the signal from the blot. After exposure, the film was placed into a tray of developing solution (Kodak) for one minute, followed by a tray of fixer solution (Kodak) for one minute, before being placed into a tray of water and then hung to dry. After drying, the protein ladder was marked on the film by overlaying the film on top of the blot using the position of the fluorescent marker.

6.3.5.3 Anti-β-actin antibody incubation

β-actin and GFAP have similar molecular masses (42kDa and 50kDa, respectively). In addition, both of these primary antibodies were raised in rabbits. Therefore it was necessary to strip the membrane of both primary and secondary antibodies.

Stripping was performed using a buffer containing 100mM β-mercaptoethanol, 62.5mM Tris-HCl (pH 6.7), and 2% SDS. This was mixed under the fume hood, and decanted into an airtight container with the membranes and placed in a shaking water bath pre-heated at 60°C, for 30 minutes. After stripping, the solution was poured away under the fume hood, and the membranes were washed in PBS Tween using large volumes of wash buffer, twice for 10 minutes each. This is to ensure removal of the β-mercaptoethanol.

Incubation with the polyclonal rabbit anti-actin primary antibody (Santa Cruz C11 sc-1615-R) was performed using the same incubation times and washing steps as for GFAP previously. However, the β-actin primary antibody was used at 1:1000 in blocking solution. After incubation with the same HRP-tagged secondary antibody as was used for GFAP, the membranes were subjected to ECL reagent and film exposure, as previously.

6.3.6 NuPAGE SDS-PAGE

The XCell SureLock Mini-Cell Electrophoresis System (Life Technologies) was used to perform Western blots, in conjunction with NuPAGE Novex 4-12% Bis-Tris 1.5mm, 10 well gels (Life Technologies). Samples were heated to 70°C for 10 minutes, after which reducing agent was added (as in manufacturer's protocol). Gels were run using MOPS buffer with antioxidant (Life Technologies). Samples were run with the Spectra Multicolor Broad Range protein ladder (Thermo Scientific), which shows apparent molecular weights of 225, 115, 80, 65, 50, 35, 30, 25, 15 and 10 kDa when run on this gel system in MOPS buffer. Gels were run for 50 minutes at 200V, according to the manufacturer's suggestion.

6.3.6.1 Blotting to membrane

Blotting was performed using the XCell II Blot Module (Life Technologies) and self-made NuPAGE transfer buffer, consisting of 25mM bicine, 25mM bis-tris, 1mM EDTA (Sigma-Aldrich), 0.05mM chlorobutanol (Sigma-Aldrich), with 10% methanol (Fisher Scientific). 20% methanol was used when running two gels in the same tank, according to the manufacturer's suggestion, to enable sufficient transfer of protein to both membranes. Amersham Hybond-P PVDF membrane (GE Healthcare) was pre-wetted in methanol for 10 seconds, washed in deionised water for five minutes. Phosphate buffered saline (PBS) was made by dissolving one PBS tablet (Sigma-Aldrich) in 200ml of deionised water. The PBS composition was 0.01M phosphate buffer, 0.0027M potassium chloride and 0.137M sodium chloride at pH 7.4. Membranes were equilibrated in PBS with 0.1% Tween 20 (Sigma-Aldrich) for 10 minutes prior to transfer. The blot module was set up such that the wetted membrane was placed on the anode side of the gel, surrounded by 3MM filter paper (Whatman) and blotting pads, all saturated in transfer buffer. The blot module was filled with NuPAGE transfer buffer containing antioxidant. Blots were run at 30V for one hour, as suggested by the manufacturer. After transfer, membranes were placed between 3MM filter paper (Whatman), wrapped in Saran wrap, and stored at 4°C overnight.

6.3.6.2 Antibody incubations

GFAP antibody incubation was performed as previously when using self-cast SDS-PAGE gels. However, the volume of polyclonal rabbit anti-GFAP primary antibody (Dako Z0334) in blocking solution was increased from 2ml to 5ml to ensure complete saturation of the membrane.

Membrane stripping was also performed as previously, with the incubation time reduced to five minutes. β-actin antibody incubation, ECL detection, and film exposure were performed identically as previously described.

6.3.6.3 Quantification of GFAP

Quantification was performed using image processing software, ImageJ (NIH). The tool allows for quantification of gels. This is achieved by analysing rectangular selections of the image to allow analysis lane by lane. The intensity of the bands are then converted to a graph, displaying peaks of intensity. By using the straight line tool, a line is drawn across the base of the peaks, and the wand tool is used to measure the area between the peaks and the straight line. The area measured on the GFAP gel image was then divided by the area measured from the β-actin image, for each sample. This allowed GFAP levels for each sample to be normalised to the β-actin control.

6.4 Results

RNAseq and qPCR are well established techniques for quantification of mRNA. However, for protein-coding genes, it is the protein which plays the active functional role, not the RNA. Since RNA expression does not always correlate with protein levels, understanding protein expression and localisation is vital in ameliorating the physiological changes that occur in the hypothalamus in *Gnasxl*$^{m+/p-}$ mice. Therefore, after identifying a decrease in *Gfap* by RNAseq and qPCR analyses, it was important to investigate whether the same was true of GFAP protein and how its expression and localisation is affected in *Gnasxl*$^{m+/p-}$ mice.

6.4.1 Using Western blots to quantify GFAP

Western blots serve as a useful method for relative quantification of protein. Using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels, protein lysates are electrophoresed in order to separate protein by molecular weight. Protein samples are loaded into wells at the top of the gel with a loading buffer, and an electric current causes the protein to migrate through the gel. Much like agarose gel electrophoresis. After electrophoresis, the protein is transferred to a membrane that can be incubated with antibodies for the purpose of quantification of specific proteins.

There are now improved systems in which to conduct western blots. One of these is the XCell II Blot Module (Life Technologies), which uses pre-cast NuPAGE polyacrylamide gels (Life Technologies) containing Bis-Tris base. These pre-cast gels have a better reproducibility, since every batch is made to the same standard. When preparing standard SDS-PAGE gels, there is a greater chance of error which causes the result to be more variable. The NuPAGE gel system has the advantage of preventing protein degradation, and produces sharper protein bands, which is necessary for accurate quantification.

Thus, protein lysates were next run on the NuPAGE system. A reducing agent was used this time to be able to better quantify the amount of GFAP as a whole. The amount of protein loaded was reduced from 10µg to 5µg, because otherwise the intensity of the bands for the monomeric proteins would be too high to detect on X-ray film.

In order to ensure low technical variability and high reducibility, duplicates of four samples were run on a NuPAGE 4-12% Bis-Tris Gel of 1.5mm thickness with 10 wells (Life Technologies) using MOPS SDS running buffer (Life Technologies). Instead of having a stacking gel on top of the resolving gel, this gel has a linear gradient of acrylamide, from 4% at the top of the gel, to 12% at the bottom of the gel. This allows for the greatest separation of mid-sized proteins, such as GFAP and β-actin which have molecular masses of 50kDa and 42kDa respectively.

The results show a varied amount of GFAP between all samples, both for wildtype and *Gnasxl*$^{m+/p-}$ samples (Figure 6.5). In addition, despite using the same wildtype samples as previously, distinctly less GFAP was detected for the 152.1 wildtype sample than for the 152.2 sample. Results from the self-cast gel had suggested that there was a similar amount of GFAP between these two wildtype samples. Levels of β-actin were shown to be similar across all samples (Figure 6.6). The comparable amount of β-actin between the wildtype samples suggest that this was not due to unequal loading (Figure 6.7).

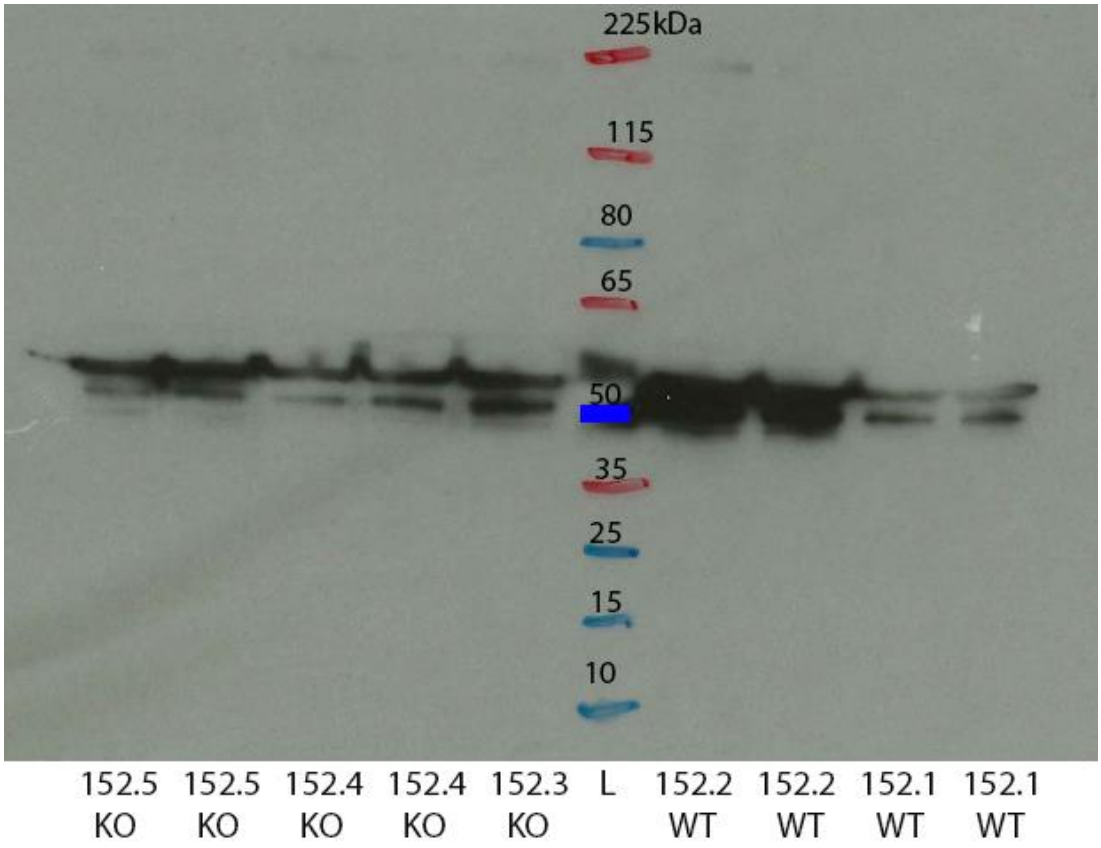

Figure 6.5: GFAP levels identified in NuPAGE Western blot of protein lysates from wildtype and *Gnasxl*$^{m+/p-}$ mice

X-ray image of the Western blot performed on the NuPAGE gel system after incubation with anti-GFAP antibody alongside the Fermentas #SM1841 ladder (L). Wildtype = WT; *Gnasxl*$^{m+/p-}$ = KO. Due to a different gel type and electrophoresis buffer, the protein band sizes migrate at different paces to the previous gel. Apparent protein band sizes are however provided by the manufacturer for various gel and buffer combinations, which are displayed in the figure.

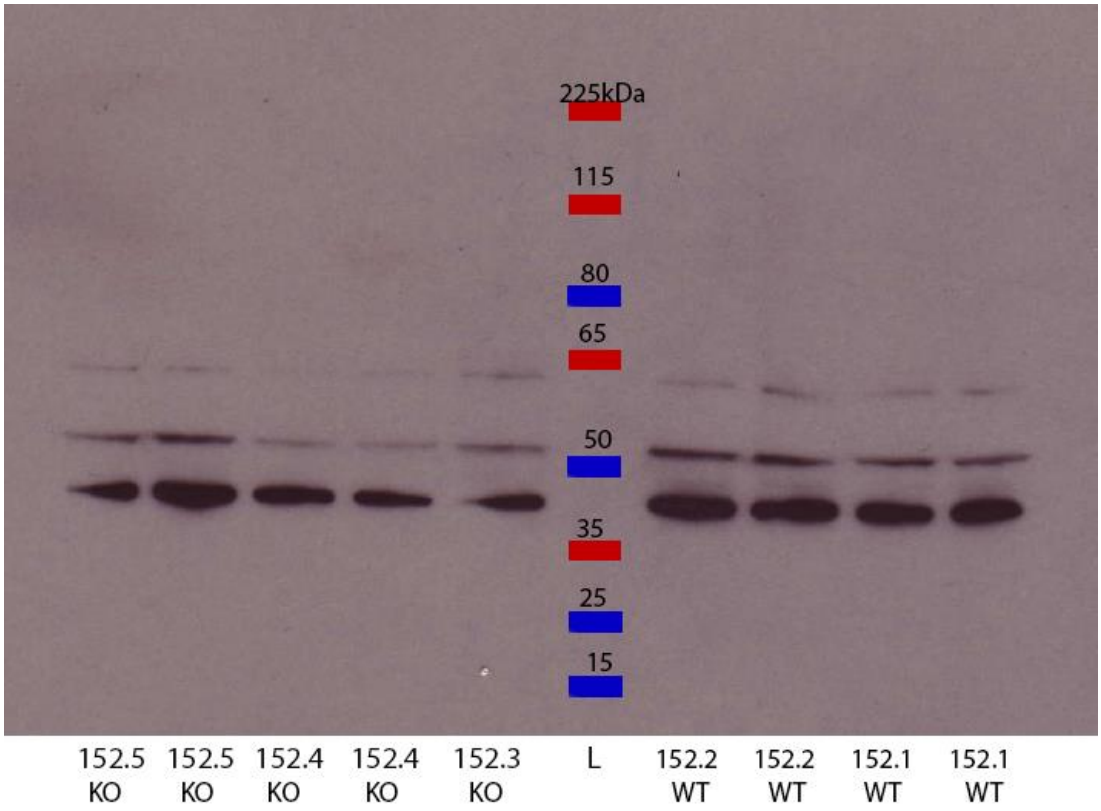

Figure 6.6: β-actin levels identified in NuPAGE Western blot of protein lysates from wildtype and *Gnasxl*$^{m+/p-}$ mice

X-ray image of the Western blot performed on the NuPAGE gel system. Wildtype = WT; *Gnasxl*$^{m+/p-}$ = KO. The membrane was stripped of anti-GFAP antibodies using β-mercaptoethanol and incubated with anti β-actin antibody. This was run alongside the Fermentas #SM1841 ladder (L), and apparent protein sizes are indicated.

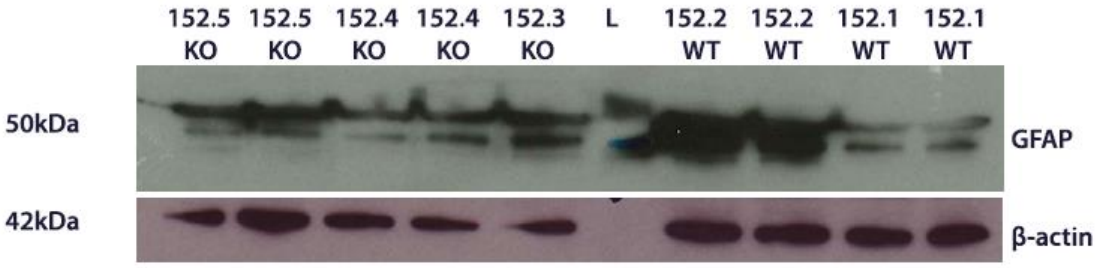

Figure 6.7: Comparison of GFAP and β-actin levels observed in the NuPAGE Western blot

In addition, to ensure a greater statistical power to detect smaller effect sizes, a larger number of samples was used across two separate gels. These gels were loaded, electrophoresed, and blotted together to reduce error and bias between samples loaded on the different gels. In total, hypothalamic lysates from six wildtype mice and eight *Gnasxl*$^{m+/p-}$ mice were run.

While the results showed a strong signal for GFAP across most samples, three wildtype samples that had not been used in the previous Western blot analyses demonstrated strikingly low levels of GFAP (Figure 6.8). Crucially, the levels of β-actin were observed to be equal across samples.

In order to quantify the relative levels of GFAP expression, a graphical analysis program, ImageJ was used. This tool has a specific function for analysing gel images, and creates an intensity graph for each lane. From this, it is possible to draw a line across the base of the curve, and use the area underneath the curve as a relative measure of protein levels. In order to normalise the samples to the levels of β-actin in each sample, the relative levels of GFAP were divided by the relative level of β-actin for each sample. A student's t test on all samples showed a significantly greater level of GFAP in *Gnasxl*$^{m+/p-}$ hypothalami ($P<0.0001$). A significant increase was still observed when excluding the three wildtype samples 123.2, 115.1, and 104.2 ($P<0.01$). However, due to the variability of results observed with the Western blot technique on these samples, it would be advisory for further Western blots to be performed. This would identify whether the result observed for wildtype samples 123.2, 115.1, and 104.2 were truly low. Furthermore, if these results are excluded on the basis of technical error, only three wildtype samples remain for analysis, which does not provide a sound statistical base for analysis. It would also be necessary to have more hypothalamic lysates available for analysis, but samples were restricted due to the high perinatal mortality of the mice, and the use of the tissues for other purposes, such as sectioning and RNA extraction.

Because GFAP is not expressed universally across the hypothalamus, accurate dissection of the hypothalamus is therefore important as it could result in variation of GFAP protein levels in whole hypothalamic lysates.

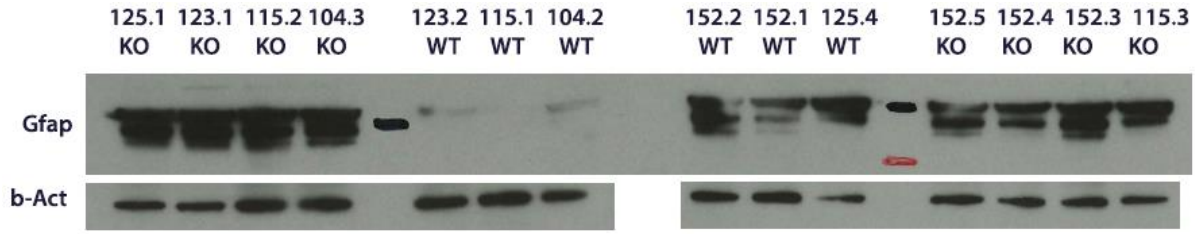

Figure 6.8: Combined X-ray images from the NuPAGE Western blots using all available lysate samples

Performed on eight *Gnasxl*$^{m+/p-}$ hypothalamic protein lysates (KO), and six wildtype (WT) samples. Three wildtype samples, 123.2, 115.1, and 104.2 show extremely low GFAP levels.

Sample	GFAP	β-actin	GFAP/β-actin
125.1 KO	21516	6597	3.26
123.1 KO	19771	6479	3.05
115.2 KO	18513	8988	2.06
104.3 KO	23233	8875	2.62
123.2 WT	2267	8729	0.26
115.1 WT	593	9217	0.06
104.2 WT	1662	6128	0.27
152.2 WT	14014	7043	1.99
152.1 WT	8311	8086	1.03
125.4 WT	10204	4530	2.25
152.5 KO	18764	7288	2.57
152.4 KO	16770	6427	2.61
152.3 KO	23199	6799	3.41
115.3 KO	16622	5933	2.80

Table 6.1: Relative quantification of protein content observed from ImageJ analysis of Figure 6.8

Numbers represent the areas underneath the curve – relative to the intensity of the gel bands. The final column shows the normalised expression of GFAP relative to that of the control β-actin. It is these normalised figures that were used for t test analysis.

6.4.2 Localisation of GFAP in adult mice

The hypothalamus contains both glial cells and neurons. In addition, the hypothalamus is a complex region of the brain with many subregions with various distinct functions. Therefore, although there were conflicting results through Western blot analyses, it may be that changes in GFAP levels could be present, but might instead be localised to particular regions.

Given its functions as an intermediate filament and as a marker of glial cells, it was important to investigate whether any changes in GFAP protein levels might be due to fewer *Gfap*-expressing cells in the hypothalamus, or due to a decreased expression of GFAP within cells. Firstly, we wanted to investigate whether there were any differences in the localisation of GFAP within the hypothalamus.

To visualise the localisation and intensity of GFAP levels in the mouse hypothalamus, slides of coronal hypothalamus sections from *Gnasxl*$^{m+/p-}$ mice and wildtype mice were subjected to immunofluorescence. Slides were incubated with a rabbit anti-GFAP primary antibody (Dako Z0334). This primary antibody is able to bind to all GFAP isoforms (Kamphuis et al., 2012). Donkey anti-rabbit Alexa488-tagged secondary antibody (Jackson ImmunoResearch) was used and the fluorescence was analysed by epifluorescence microscopy. Alexa488 is a fluorescent dye that absorbs light at wavelength 495nm and emits at 519nm. This was counter-stained with DAPI, a fluorescent compound that is able to bind to the minor groove within DNA. This emits light at 461nm and results in staining of the nuclei of cells. It is used as a control, but also to be able to visualise the structures within the tissue. In particular, in the hypothalamus, densely packed nuclei such as the SCN and the PVN are easily identified by the high density of cells.

6.4.2.1 Ependymal cells

Ependymal cells are a type of glial cell which line the edge of the brain surfaces. They line the surface that is in contact with cerebrospinal fluid, which they are able to produce, and absorb using microvilli on their surface. The ependymal cells show a high expression of GFAP in both wildtype mice and *Gnasxl*$^{m+/p-}$ mice (Figure 6.9). This appears as a bright green fluorescence coating the peripheral layer of the brain. Intensity of fluorescence against GFAP appears to be no different for ependymal cells between wildtype and *Gnasxl*$^{m+/p-}$ mice.

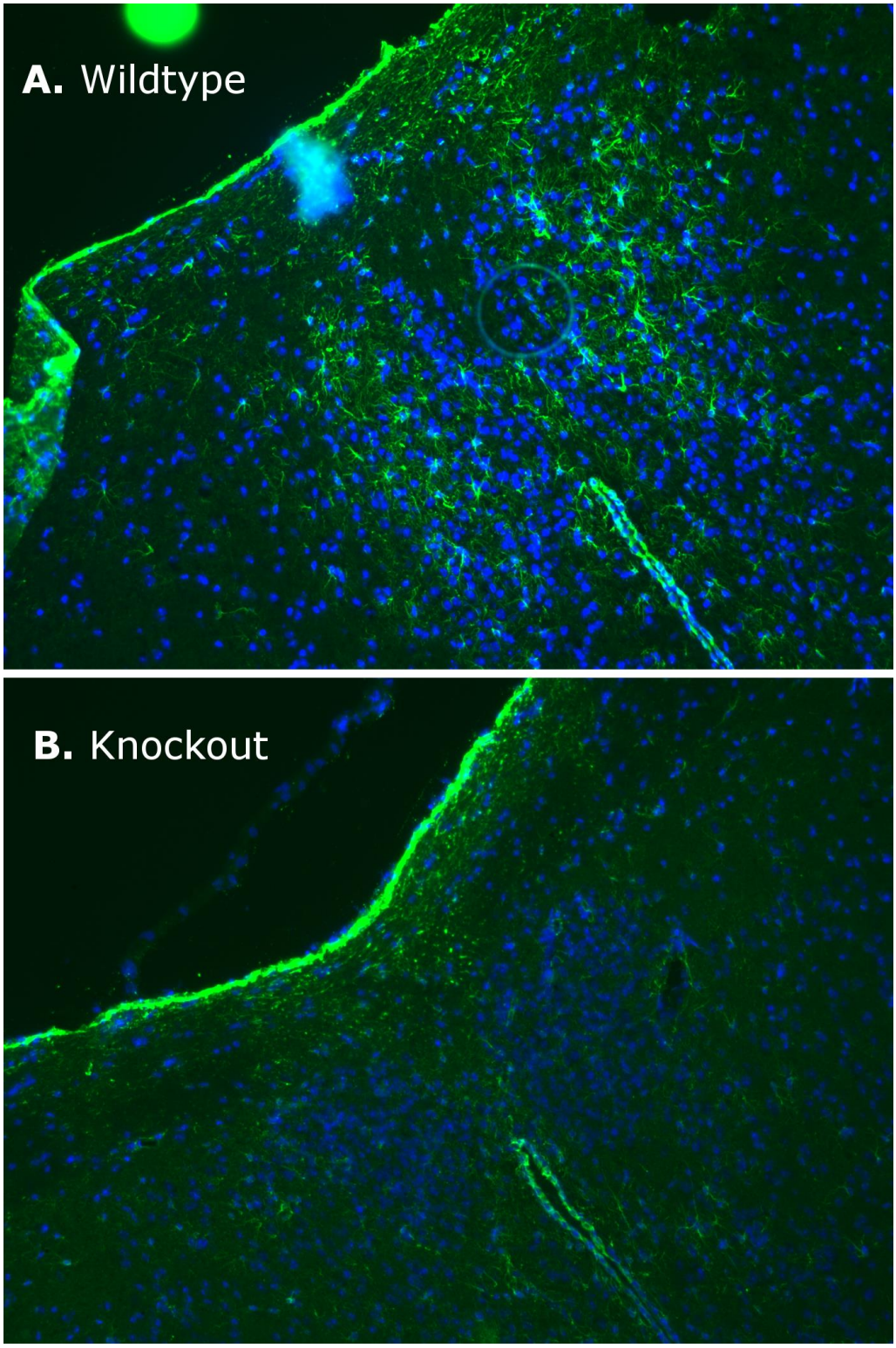

Figure 6.9: Immunofluorescence of the ventral side of the hypothalamus at the SCN of a wildtype and knockout Gnasxl$^{m+/p-}$ mice

Coronal sections of mice hypothalami at -0.70mm bregma at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue). **A.** Wildtype **B.** Gnasxl$^{m+/p-}$. Ependymal cells, around the periphery of the brain surface widely express GFAP.

6.4.2.2 Tanycytes

The hypothalamus is characterised by the third ventricle, which is filled with cerebrospinal fluid and runs through the core of the hypothalamus. The third ventricle is lined with tanycyte cells, which are characterised by their elongated morphology, with extensive projections directing away from the third ventricle to other regions of the hypothalamus. Towards the middle and the dorsal end of the third ventricle, α1-type tanycytes are found, which project towards the DMH and the VMH. Towards the ventral end of the third ventricle, α2 tanycytes are found in the Arc with extensive projections towards the VMH. β1 and β2 tanycytes are found in the median eminence region, on the ventral edge of the hypothalamus by the Arc.

Surrounding the middle of the third ventricle, it is possible to observe the various types of GFAP-expressing cells in wildtype brains (Figure 6.10). Tanycytes line the third ventricle, with projections directing outwards towards the DMH. In addition, it is possible to observe the star-like morphology of GFAP-expressing astrocytes further away from the third ventricle. Upon using greater magnification, the tanycyte projections can be seen more clearly emanating perpendicularly to the third ventricle lining (Figure 6.11).

In addition, astrocytes are present further laterally from the third ventricle, in the DMH region of the hypothalamus (Figure 6.10). These cells are distributed amongst the neurons within the hypothalamus. This is typical of these types of glial cells, which help to maintain the chemical environment and modulate synaptic functions of neurons.

In the *Gnasxl*$^{m+/p-}$ mice, there is a distinct loss of GFAP around the centre of the third ventricle at bregma -1.70mm. Although GFAP-expressing cells are present, the expression levels are very much decreased (Figure 6.12). Using a greater exposure time, it is possible to identify the tanycytes surrounding the third ventricle in *Gnasxl*$^{m+/p-}$ mice (Figure 6.12). The structures of the tanycytes are very similar to those seen in wildtype mice, suggesting that it is representative of an underexpression of GFAP (Figure 6.11, arrowhead). In addition, the tanycytes still project towards the DMH, although there is a clear reduction in GFAP immunofluorescence from these cells (Figure 6.12). Towards the DMH, there is a distinct loss of the astrocytes that characterised this region in the wildtype animals (Figure 6.10, arrowhead).

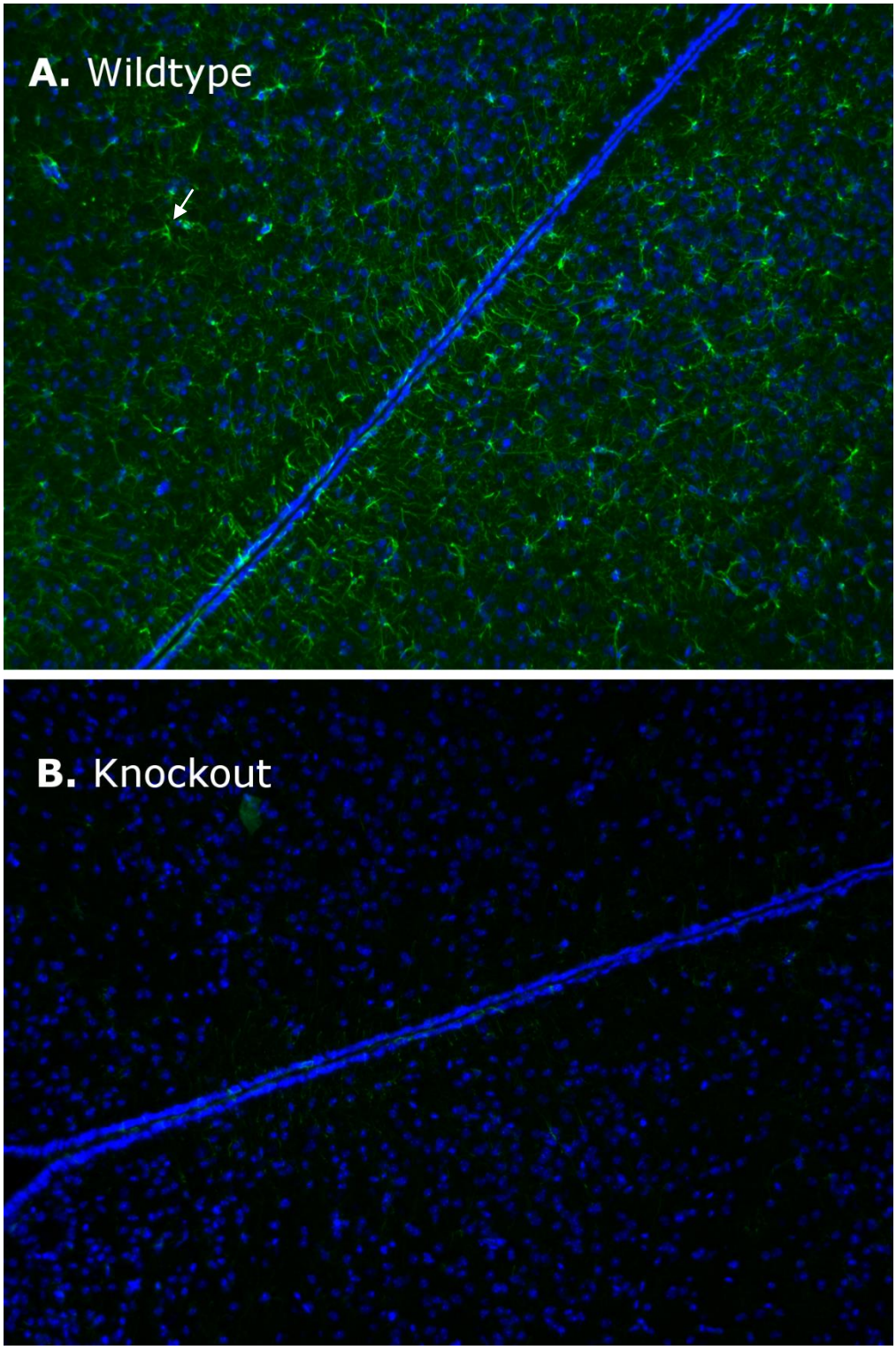

Figure 6.10: Immunofluorescence of the 3V of wildtype and Gnasxl$^{m+/p-}$ mice

Coronal sections of a wildtype mouse and Gnasxl$^{m+/p-}$ mouse brain at bregma -1.70mm at 10X magnification. The ventral side of the brain is located to the bottom left of the photographs. Sections were stained for GFAP (green) and DAPI (blue). **A.** Wildtype mouse. **B.** Gnasxl$^{m+/p-}$ mouse. Arrowhead: astrocytic-like cell in the DMH

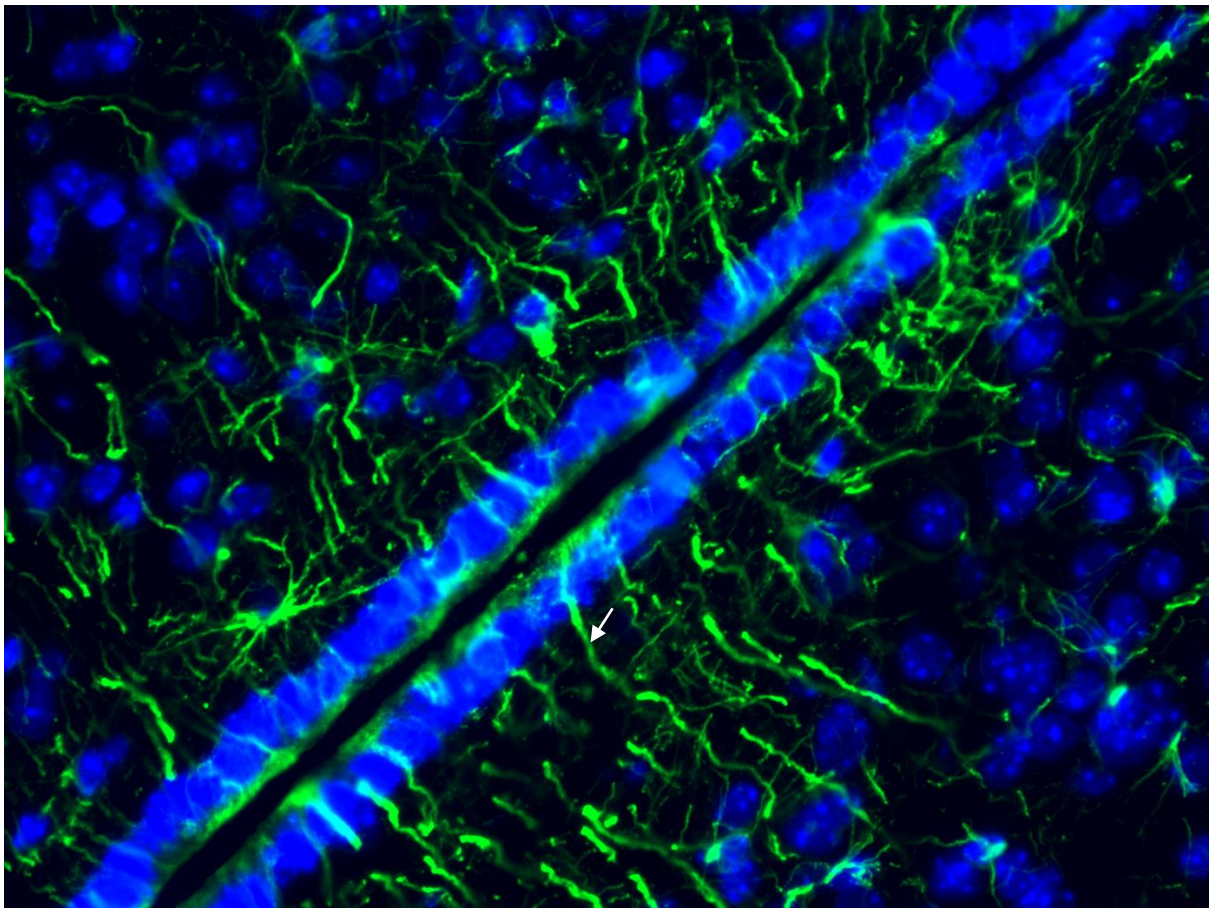

Figure 6.11: Immunofluorescence of the 3V of a wildtype mouse at greater magnification

Photograph taken at bregma -1.70mm at 40X magnification. The ventral side of the brain is located to the bottom left of the photograph. Sections were stained for GFAP (green) and DAPI (blue). Arrowhead: tanycyte projection from the wall of the 3V.

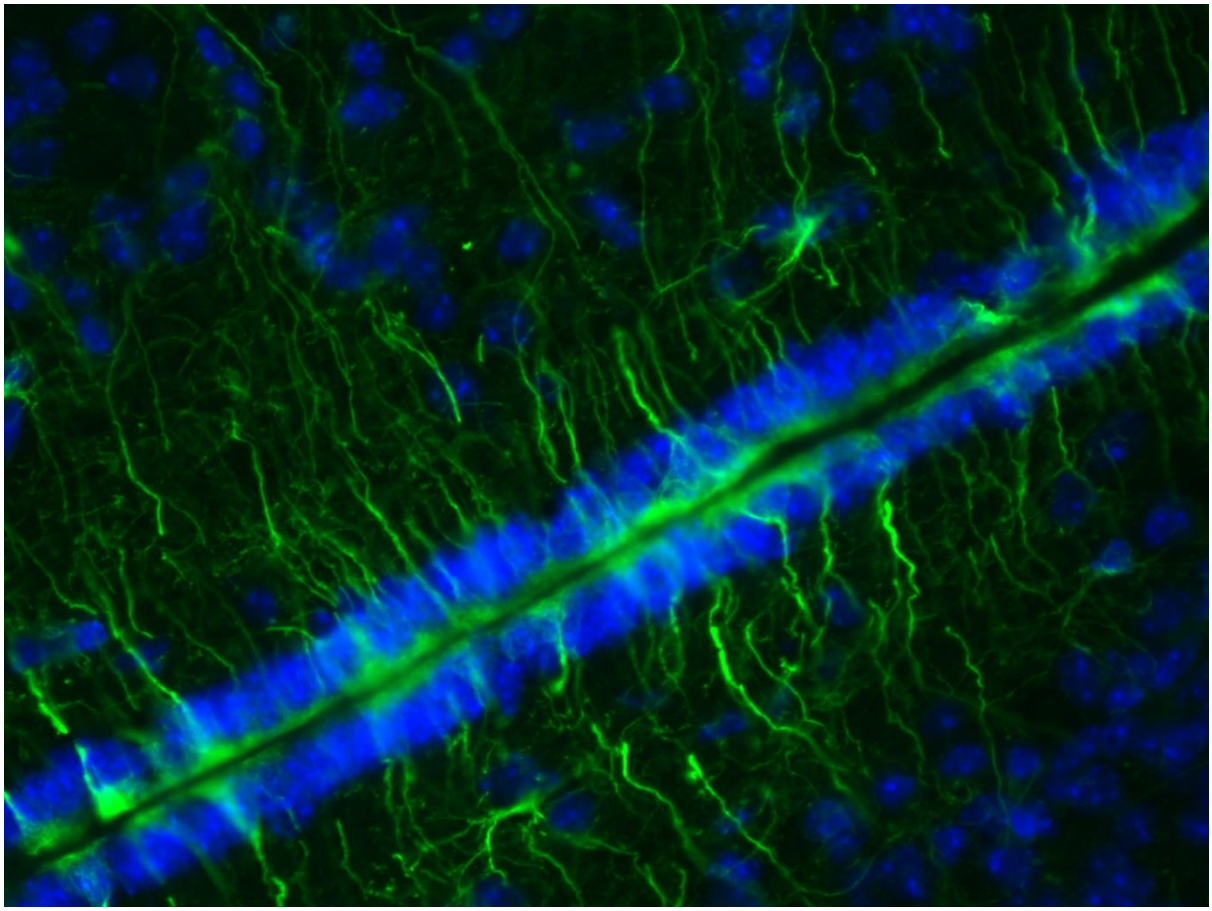

Figure 6.12: Immunofluorescence of the 3V from a coronal section of a Gnasxl$^{m+/p-}$ mouse

Section cut at bregma -1.70mm at 40X magnification. Sections were stained for GFAP (green) and DAPI (blue). Exposure time to GFAP was doubled in order to show the structures more clearly.

In one of the *Gnasxl*$^{m+/p-}$ mice, an apparent rupture of the third ventricle was identified towards the dorsal side (Figure 6.13). The lining of the third ventricle appears to have been torn, and cells expressing GFAP have migrated into the third ventricle space, likely as a response to the damage. This is typical of a glial scar formation, whereby astrocytes proliferate into regions of damage. Expression of *Gfap* is increased in the astrocytes in order to extend their processes and change their morphology to respond to the damage. The mouse was healthy when anaesthetised and killed and did not seem to cause any pathology. In addition, GFAP expression patterns in this mouse in the hypothalamus did not deviate from those seen in other *Gnasxl*$^{m+/p-}$ mice. This is the first time that such a phenomenon has been identified in *Gnasxl*$^{m+/p-}$ mice, although it is not yet known whether this is related to its genetic background.

6.4.2.3 Arcuate nucleus

GFAP levels increase towards the ventral side of the third ventricle, where it is surrounded by the Arc. At this region, the third ventricle becomes wider, allowing access to the cerebrospinal fluid by the cells of the Arc. Surrounding the third ventricle within the Arc, tanycytes send out extensive projections towards the VMH.

Adult mouse brain sections were analysed for GFAP by immunofluorescence, in order to ascertain whether there were any changes in GFAP levels and its localisation within the Arc. Three animals of wildtype and three *Gnasxl*$^{m+/p-}$ mice were analysed. The Arc is identified by the slightly denser clustering of cells around the base of the third ventricle at bregma -1.70mm. This can be observed using DAPI staining of cell nuclei to identify the region.

Wildtype adult mice show an intense expression of GFAP within the Arc at either side of the base of the third ventricle. The cell morphology is typical of tanycyte cells that populate the third ventricle wall. Within the Arc, these cells show their longest projections, which direct towards the VMH (Figure 6.14). Astrocytic-like cells can be observed just outside the Arc, within the VMH (Figure 6.14).

In addition, the median eminence region, located ventrally to the third ventricle, alongside the Arc at bregma -1.70mm, also widely expresses GFAP in wildtype mice (Figure 6.14).

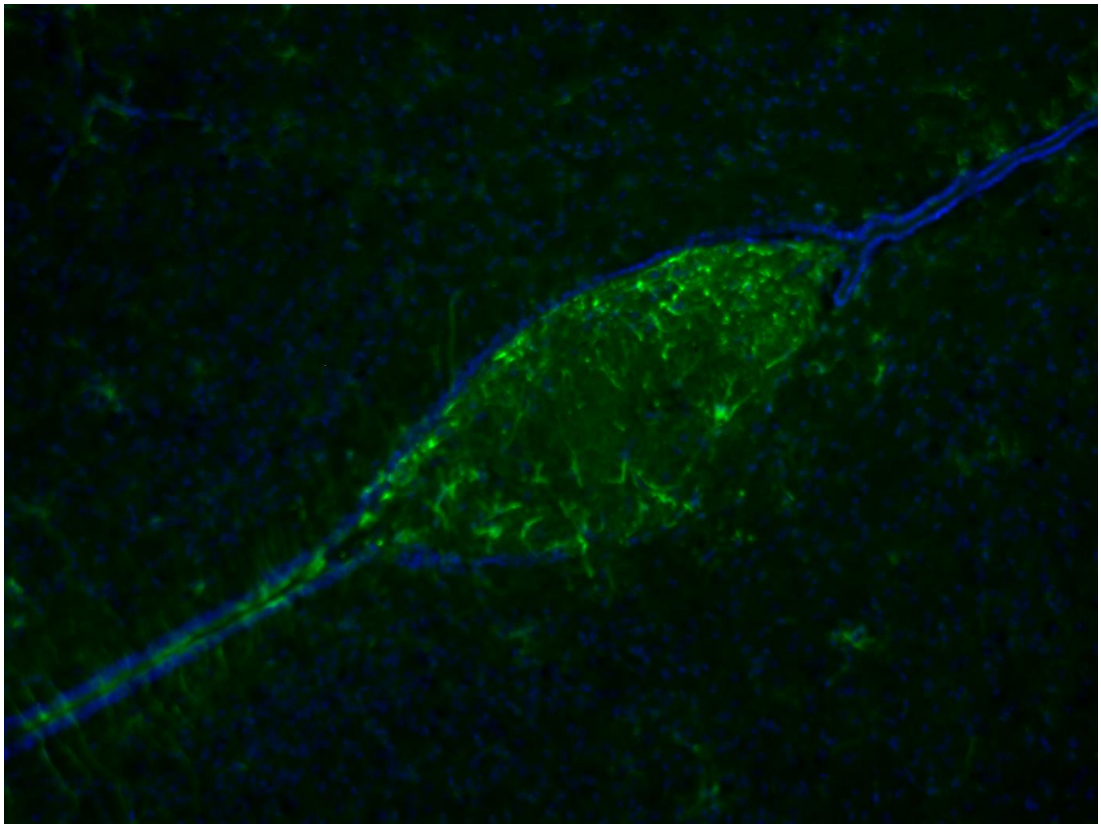

Figure 6.13: Immunofluorescence of the 3V of a Gnasxl$^{m+/p-}$ mouse showing signs of astrogliosis

Immunofluorescence of the ruptured third ventricle of the hypothalamus of a 14µm-thick coronal section of a Gnasxl$^{m+/p-}$ mouse brain at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

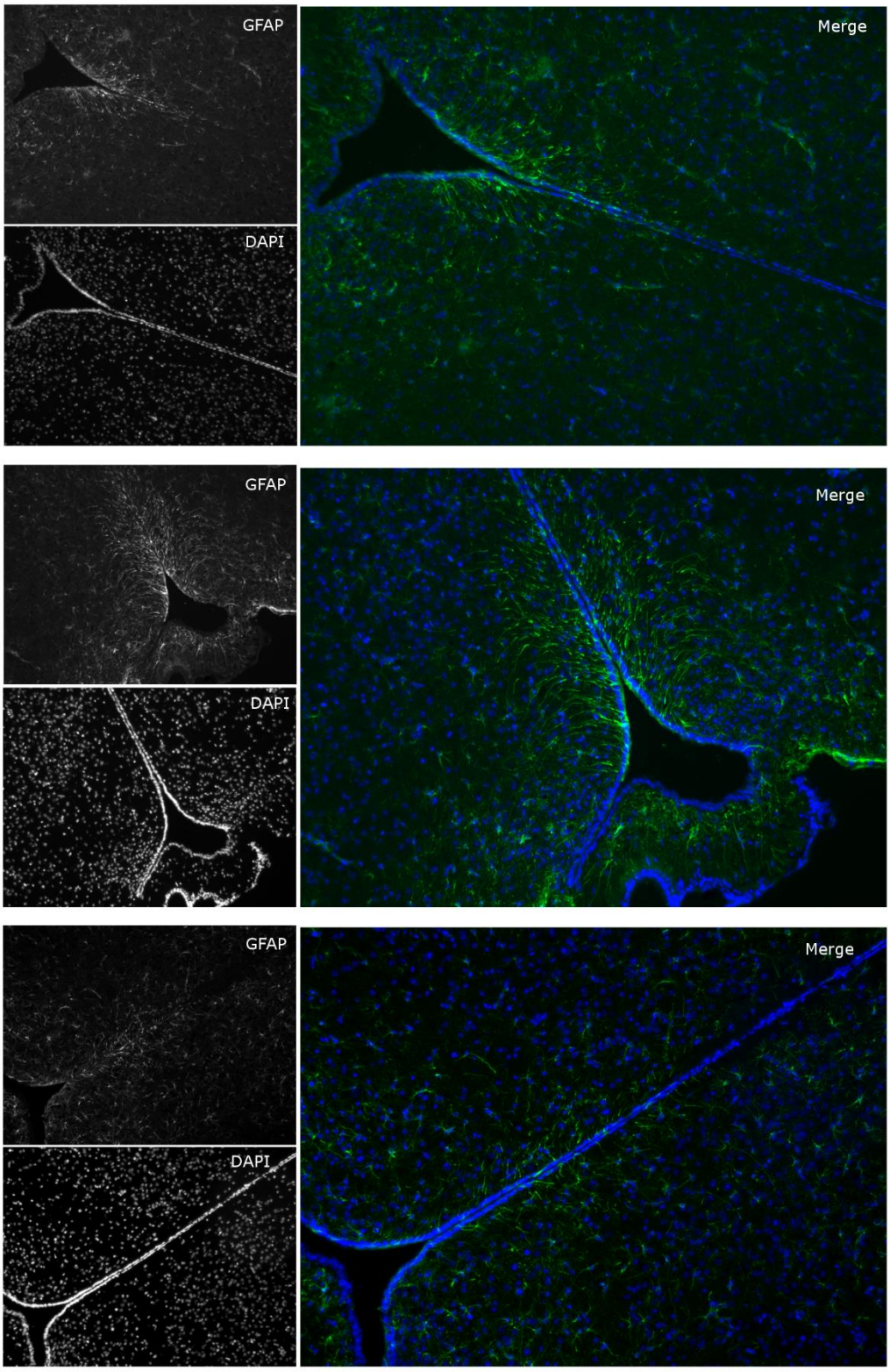

Figure 6.14: Immunofluorescence of three adult wildtype mice showing the 3V and the Arc

Immunofluorescence of the Arc of the hypothalamus of a 14µm-thick coronal sections from three wildtype mice at bregma -1.70mm 10X magnification. Sections were stained for GFAP (green) and DAPI (blue). See Figure 6.3 for schematic diagram.

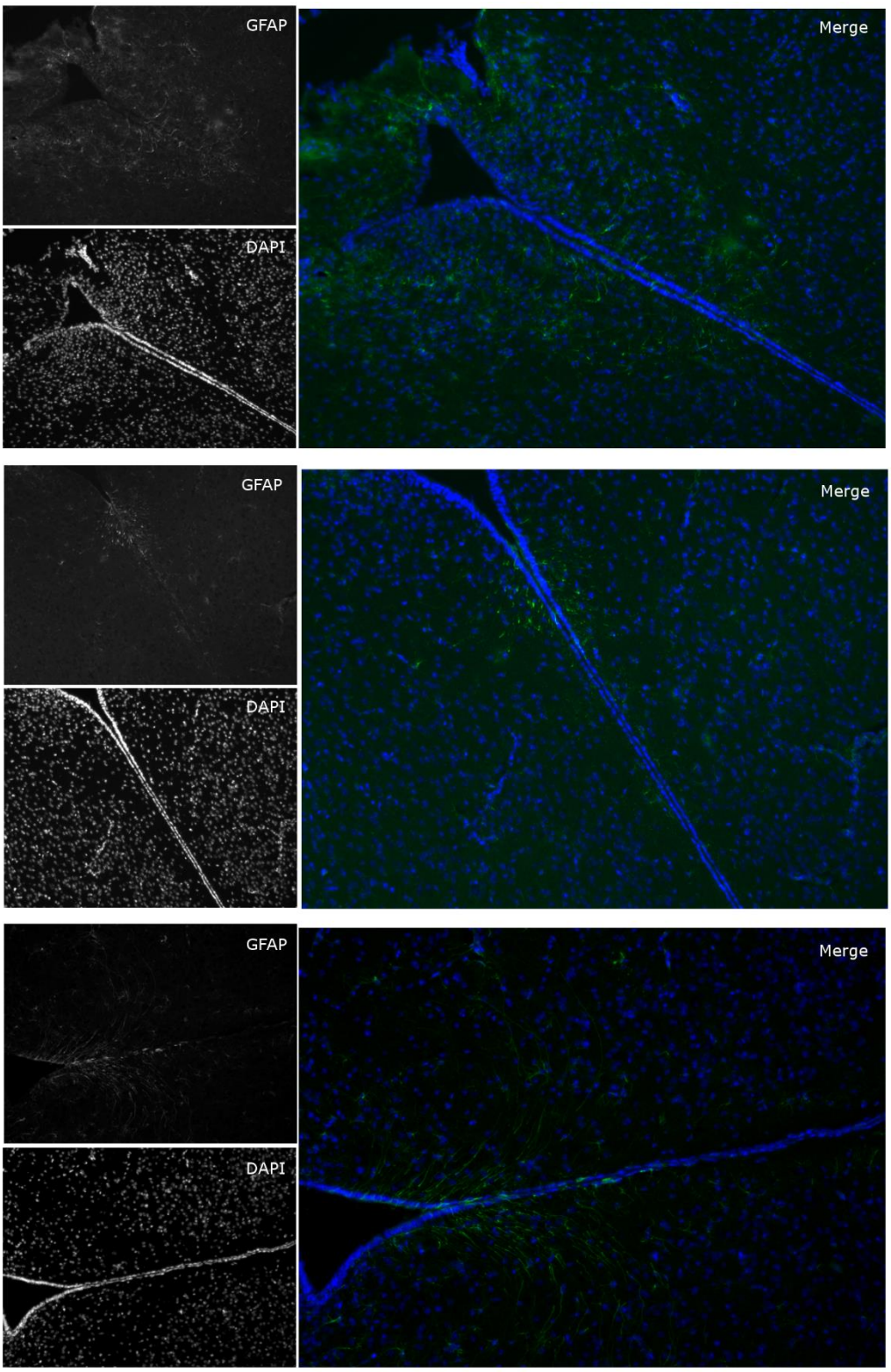

Figure 6.15: Immunofluorescence of three adult *Gnasxl*$^{m+/p-}$ mice showing the 3V and the Arc

Immunofluorescence of the Arc from three *Gnasxl*$^{m+/p-}$ mice at bregma -1.70mm at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue). See Figure 6.3 for schematic diagram.

Gnasxl$^{m+/p-}$ mice have a much reduced amount of GFAP within the Arc and surrounding regions (Figure 6.15). Although tanycyte structures show a similar morphology to wildtype, there appear to be fewer tanycyte cells expressing GFAP in this region, and tanycyte projections also appear to be reduced. However due to their long projections, it is not easy to distinguish cells from one another using immunofluorescence in order to be able to count cell numbers. Noticeably, GFAP appears to be more restricted to regions close to the third ventricle than is observed in wildtype animals, where expression of GFAP extends towards the VMH (Figure 6.14). Tanycyte projections appear shorter, and fluorescence is reduced overall in the Arc in *Gnasxl*$^{m+/p-}$ mice.

Upon increasing the exposure time, the tanycyte projections can be seen more clearly for one of the knockout mice, indicating that the cell projections do exist, but that the distribution of GFAP is less intense than in wildtype mice.

6.4.2.4 Suprachiasmatic nucleus

Anterior to the Arc, the SCN is another region of great importance within the hypothalamus. It consists of a small collection of cells adjacent to the third ventricle and is located directly above the optic chiasm. The SCN is responsible for the regulation of the circadian rhythm. Recently, it has been identified that astrocytes in the SCN play an important role in regulating the circadian rhythm.

In adult wildtype mice, immunofluorescence shows GFAP expression distributed throughout the SCN (Figure 6.16). The morphologies of the glial cells are astrocytic, which are very populous within the nuclei and are distributed evenly.

On the other hand, *Gnasxl*$^{m+/p-}$ mice show a reduced amount of GFAP within the SCN, with few astrocytes in the nuclei or within the surrounding regions. However, the ependymal cells of this region maintain their GFAP expression

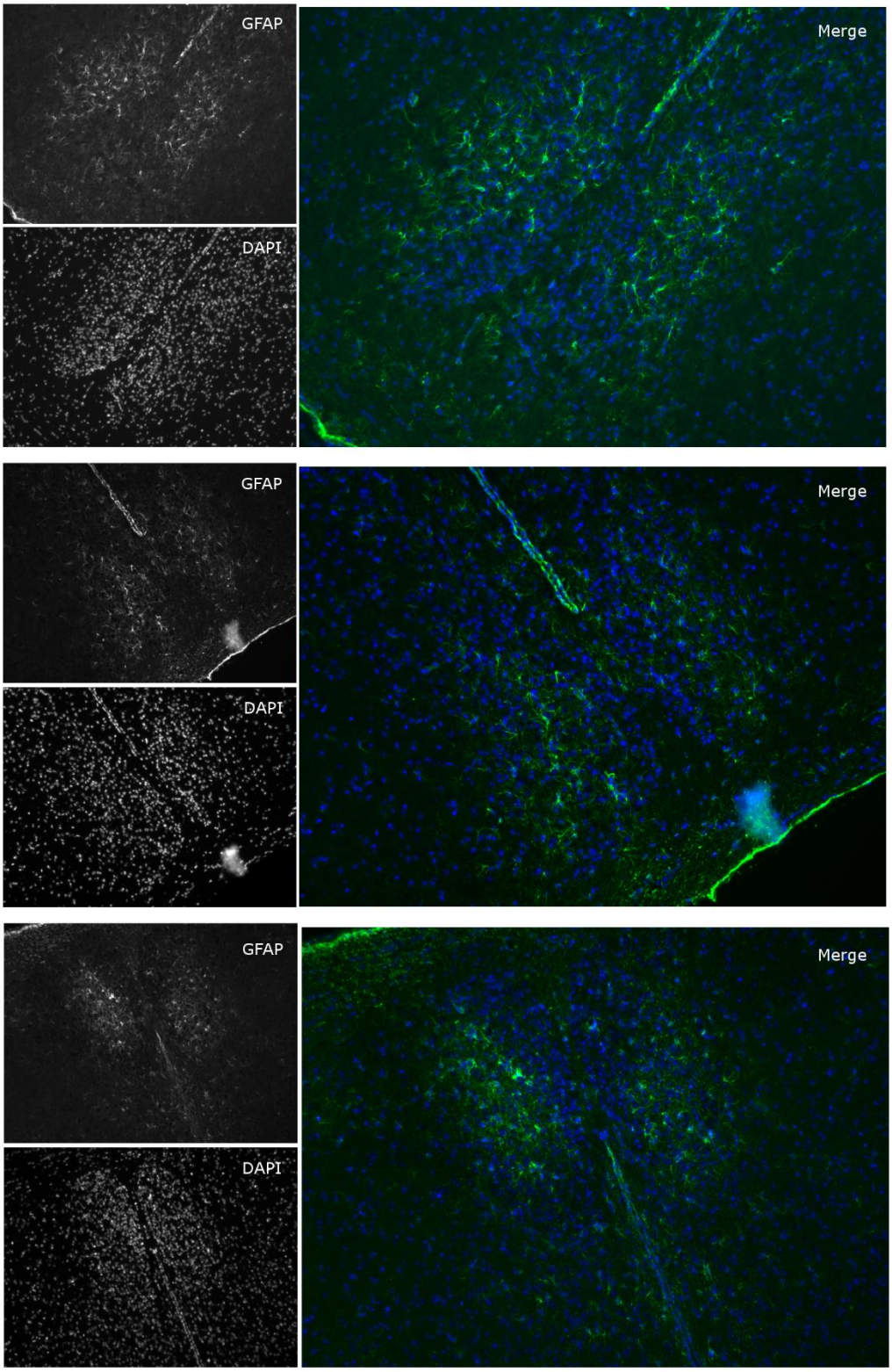

Figure 6.16: Immunofluorescence of three adult wildtype mice showing the SCN

Immunofluorescence of the SCN of the hypothalamus of a 14µm-thick coronal sections from three wildtype mice at bregma -0.70mm at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

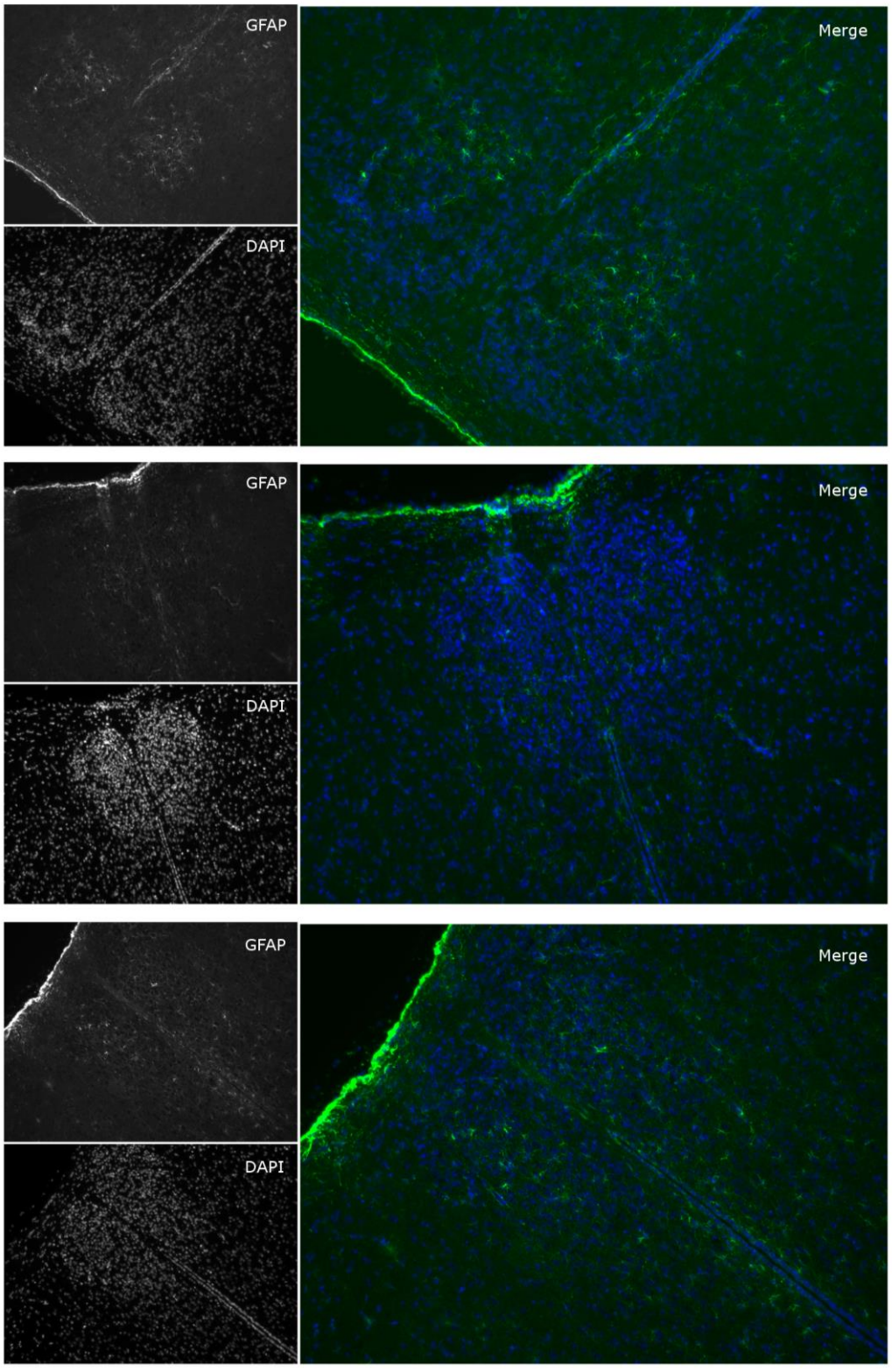

Figure 6.17: Immunofluorescence of three adult *Gnasxl*$^{m+/p-}$ mice showing the SCN Immunofluorescence of the SCN of the hypothalamus of a 14µm-thick coronal sections from three *Gnasxl*$^{m+/p-}$ mice at bregma -0.70mm at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

6.4.3 Localisation of GFAP in neonatal mice

After identifying changes in GFAP in adult mice, we were keen to understand changes during development. In particular, it was not clear whether these changes were an adaptation to chronic undernutrition, or whether there was a disruption during development that was driving these brain changes in the adults.

In order to further understand the patterns of GFAP expression in the hypothalamus, we chose to investigate how this patterning occurs during postnatal development. In particular, it was not clear whether the lack of GFAP observed in adults was due to a response to chronic undernutrition, or whether there might be a developmental problem leading to an insufficient glial population within the hypothalamus in adults. To investigate this, breeding pairs were established to create new litters for study.

Littermates were housed together and sacrificed at the same time. The first litter pair was sacrificed at postnatal day 1 (P1), and a second litter pair was sacrificed at postnatal day 5 (P5). Litter size was restricted by the high death rate of *Gnasxl*$^{m+/p-}$ mice during the early postnatal stages.

The progenitor cells of glia begin in the neural tube during embryonic development. Sonic hedgehog (SHH) is secreted from the notochord and the floor plate of the neural tube, and its antagonist bone morphogenic protein (BMP) is secreted from the roof plate. This creates a morphogen gradient that allows cells to differentiate into different types of cells depending on their location and how much SHH they receive (Ingham and Placzek, 2006). This is critical in patterning the neural tube to direct the differentiation of cells into different cell types.

Cells that express *Shh* early in development, prior to embryonic day 9.5 (E9.5), go on to produce neurons, tanycytes, and astrocytes towards the posterior regions of the hypothalamus, including the Arc and ME (Alvarez-Bolado et al., 2012). Cells expressing *Shh* after E9.5 and before E12.5 give rise to neurons and astrocytes within the tuberal central region of the hypothalamus, particularly within the VMH and DMH. After E12.5, nearly all *Shh*-expressing cells contribute hypothalamic astrocytes, representing a developmental change from neurogenesis to gliogenesis (Alvarez-Bolado et al., 2012).

Gliogenesis is a process by which the glial cells differentiate. Unlike neurogenesis, which occurs prior to birth in rodents, gliogenesis continues postnatally in the first two to three weeks of life. In humans, gliogenesis occurs during the third trimester of pregnancy.

While GFAP is only expressed in glial cells, during gliogenesis radial glia are able to divide asymmetrically to produce new neuronal precursors. During the early stages of gliogenesis, expression of GFAP is not only representative of differentiated glial cells, but also of expression within these radial glia

During development in the subgranular zone within the hippocampus has an important role in neurogenesis. In particular, radial glial cells act as a scaffold for new neurons to migrate to the granule cell layer. Radial glial cells can divide asymmetrically to produce a new stem cell, and a neuronal precursor cell. They also express GFAP, since they are a subset of astrocyte cells. Immunofluorescence of GFAP appeared very strongly in the hippocampus of wildtype mice where these radial glial cells are known to be present (Figure 6.18).

During P1, GFAP expression was observed in the wildtype hypothalamus by immunofluorescence, although at a lower level to that seen in wildtype adult mice (Figure 6.19). Since gliogenesis occurs mainly postnatally, and since pup weaning and feeding during early life are known to have important consequences on brain development, this region has not yet fully developed and GFAP-expressing cells are still developing in these regions at P1.

Limitations of sample size made it difficult to identify whether there was a difference in GFAP between wildtype and *Gnasxl*$^{m+/p-}$ mice at P1 stage. However, GFAP was observed in the Arc of *Gnasxl*$^{m+/p-}$ P1 mice. It was possible to see the tanycyte projections extending from the third ventricle (Figure 6.20), and that the projections are quite extensive (Figure 6.21). Furthermore, on closer inspection, a large collection of cells resembling astrocytes were observed towards the periphery of the hypothalamus, around the ME region (Figure 6.22). This leads to the suggestion that gliogenesis does not appear to be inhibited at the P1 stage in *Gnasxl*$^{m+/p-}$ P1 mice.

At the PVN, GFAP-expressing cells are observed stretching from the PVN across to the lateral hypothalamus close to the zona incerta (Figure 6.23). The same phenomenon is observed in *Gnasxl*$^{m+/p-}$ mice (Figure 6.24). These projections are not observed in this region in adult mice, and may be glial cells migrating towards the hypothalamus during the early stages of gliogenesis. By P5, the expression of GFAP appeared more intense, both for wildtype mice and for *Gnasxl*$^{m+/p-}$ mice (Figures 6.25, 6.26). However, there is no discernible difference between the expression observed in wildtype mice and for *Gnasxl*$^{m+/p-}$ mice in the PVN at P5.

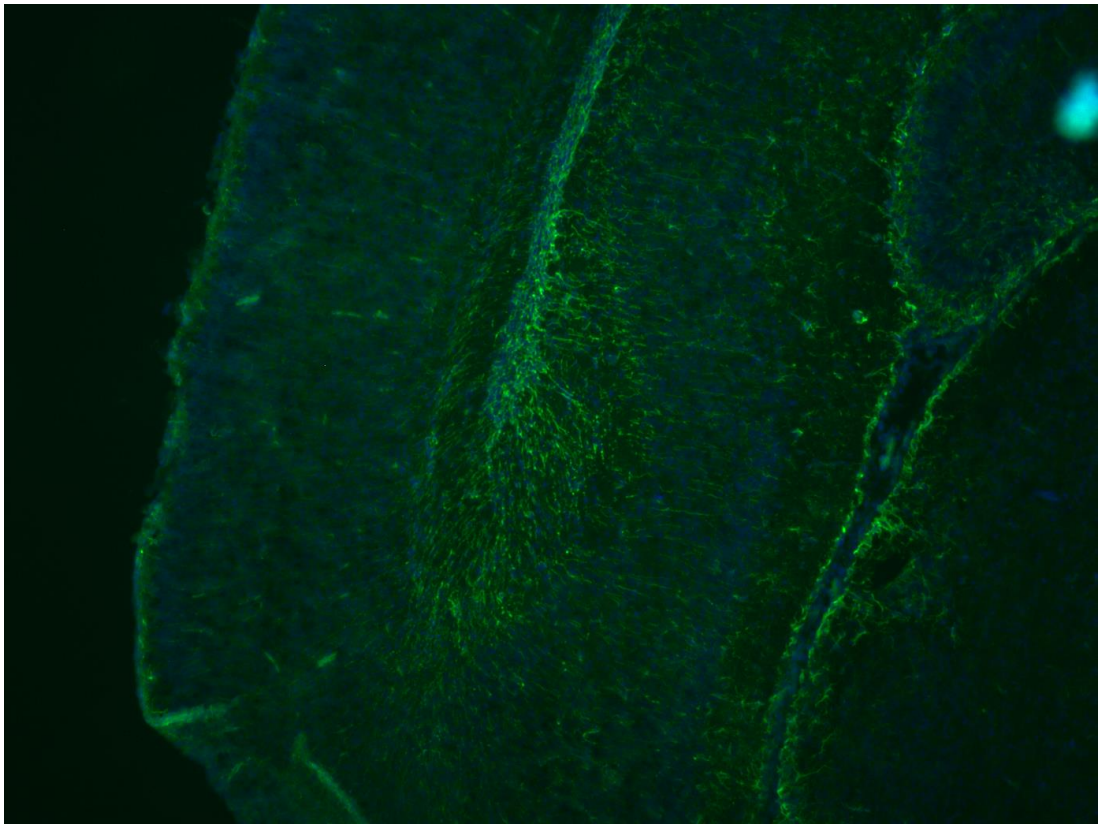

Figure 6.18: Immunofluorescence of the hippocampus of a P1 wildtype mouse

Immunofluorescence of the hippocampus of a 14µm-thick coronal sections of a wildtype mouse at P1 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

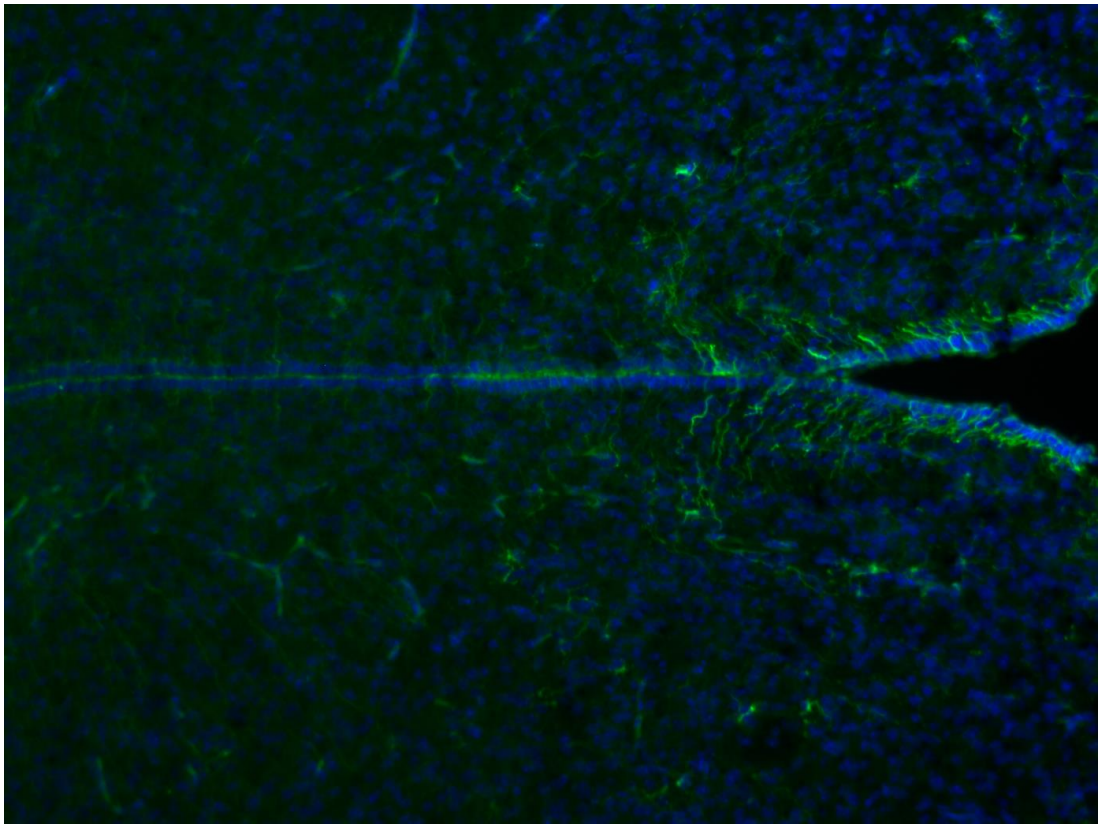

Figure 6.19: Immunofluorescence of the Arc of a P1 wildtype mouse

Immunofluorescence of the arcuate nucleus of the hypothalamus of a 14µm-thick coronal section of a wildtype mouse at P1 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

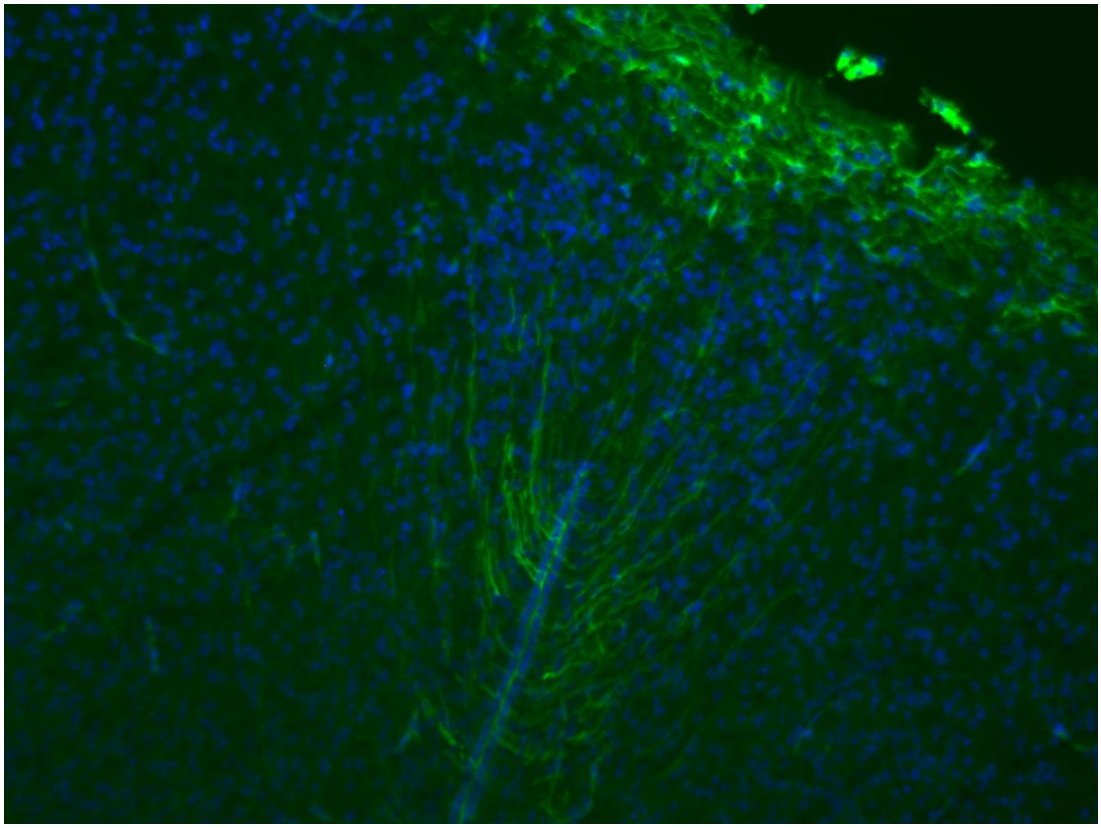

Figure 6.20: Immunofluorescence of the Arc of a P1 *Gnasxl*$^{m+/p-}$ mouse

Immunofluorescence of the arcuate nucleus of the hypothalamus of a 14µm-thick coronal section of a *Gnasxl*$^{m+/p-}$ mouse at P1 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

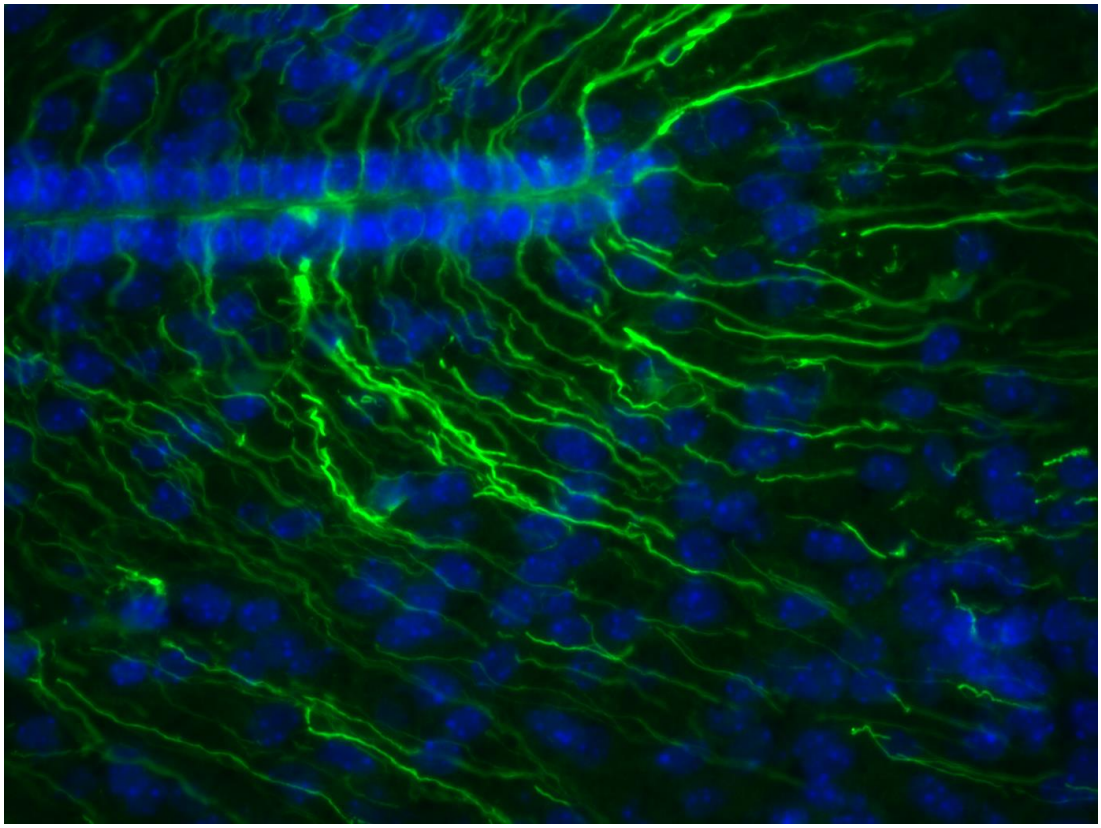

Figure 6.21: Immunofluorescence of the Arc of a P1 *Gnasxl*$^{m+/p-}$ mouse at greater magnification

Immunofluorescence of the arcuate nucleus of the hypothalamus of a 14μm-thick coronal section of a *Gnasxl*$^{m+/p-}$ mouse at P1 at 40X magnification. Sections were stained for GFAP (green) and DAPI (blue).

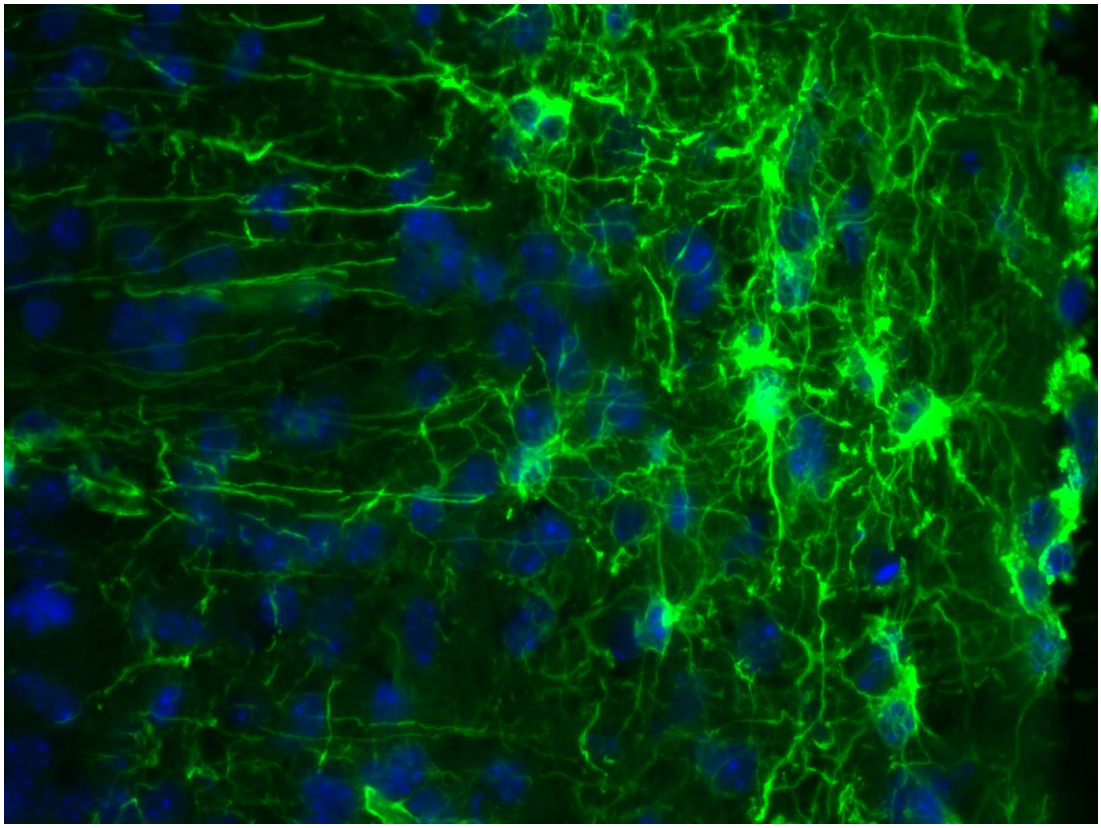

Figure 6.22: Immunofluorescence of the ME of a P1 *Gnasxl*$^{m+/p-}$ mouse

Immunofluorescence of the median eminence of the hypothalamus of a 14µm-thick coronal section of a *Gnasxl*$^{m+/p-}$ mouse at P1 at 40X magnification. Sections were stained for GFAP (green) and DAPI (blue).

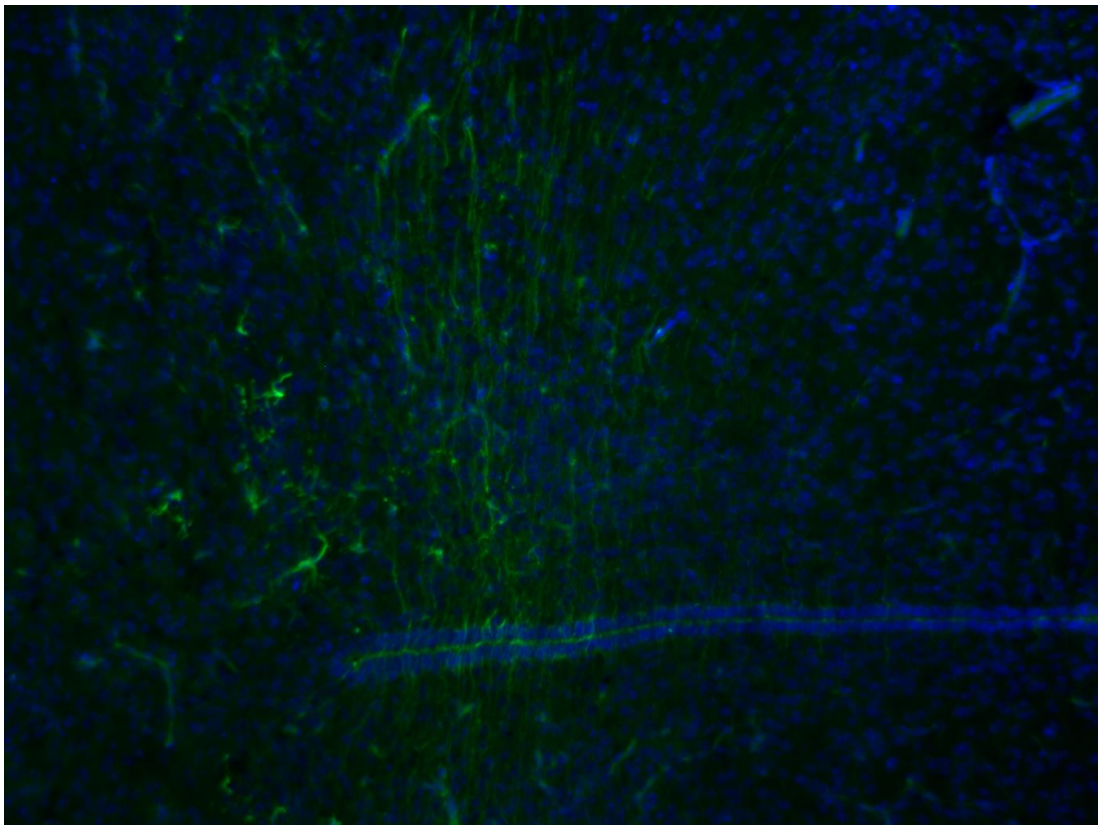

Figure 6.23: Immunofluorescence of the PVN of a P1 wildtype mouse

Immunofluorescence of the paraventricular nucleus of the hypothalamus of a 14µm-thick coronal section of a wildtype mouse at P1 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

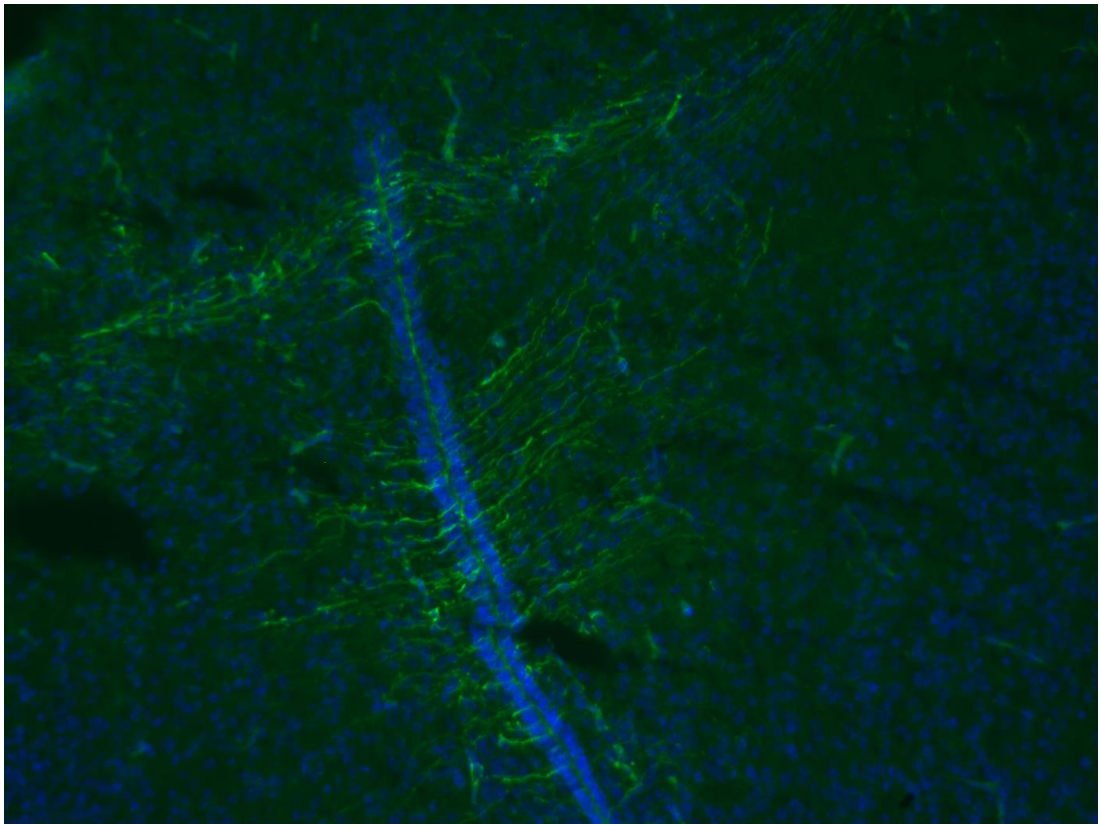

Figure 6.24: Immunofluorescence of the PVN of a P1 *Gnasxl*$^{m+/p-}$ mouse Immunofluorescence of the paraventricular nucleus of the hypothalamus of a 14µm-thick coronal section of a Gnasxl$^{m+/p-}$ mouse at P1 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

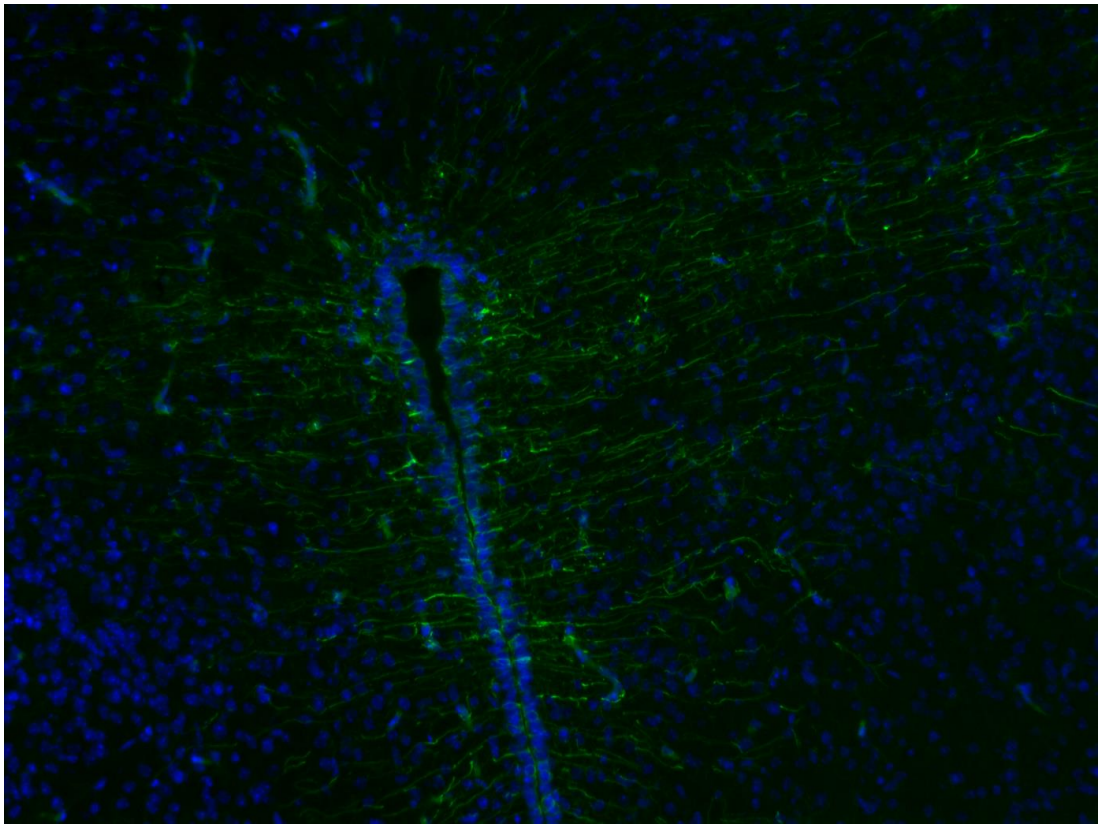

Figure 6.25: Immunofluorescence of the PVN of a P5 wildtype mouse Immunofluorescence of the paraventricular nucleus of the hypothalamus of a 14µm-thick coronal section of a wildtype mouse at P5 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

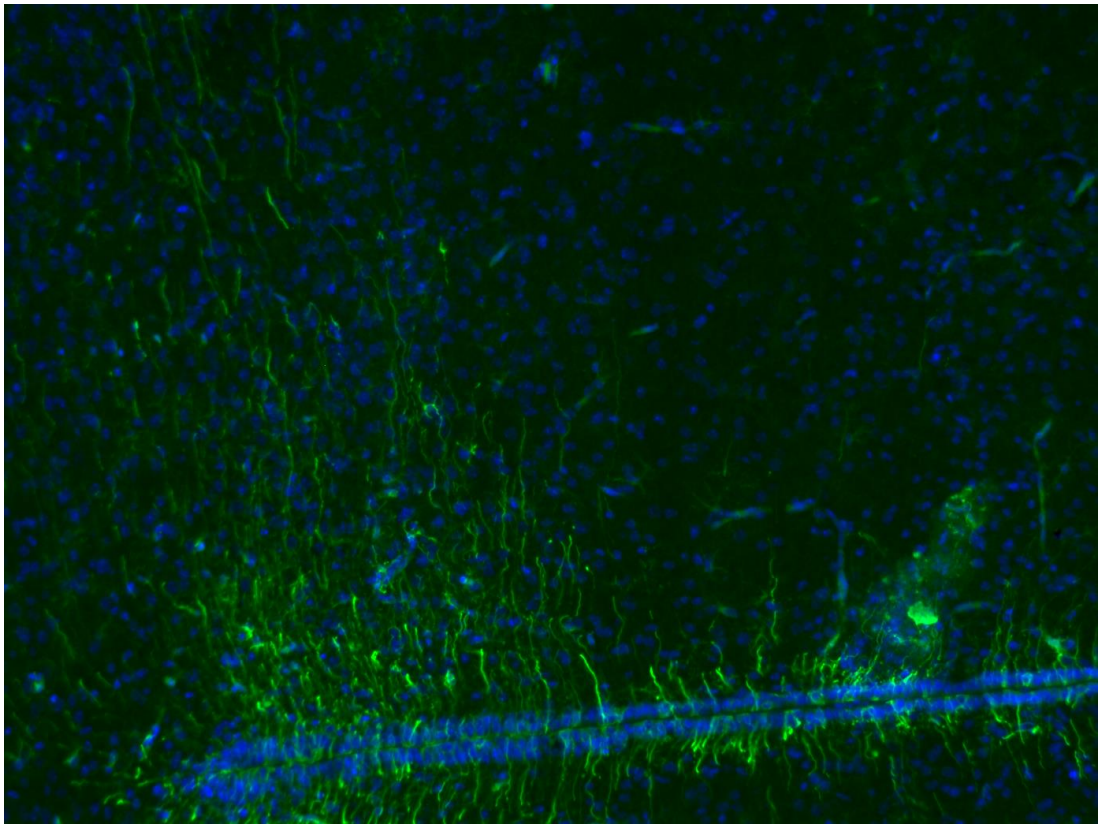

Figure 6.26: Immunofluorescence of the PVN of a P5 *Gnasxl*$^{m+/p-}$ mouse Immunofluorescence of the paraventricular nucleus of the hypothalamus of a 14µm-thick coronal section of a *Gnasxl*$^{m+/p-}$ mouse at P5 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

Expression of GFAP in the Arc of P5 mice was limited in both wildtype and *Gnasxl*$^{m+/p-}$ mice (Figures 6.27 and 6.28). Expression surrounding the third ventricle was very much limited to the Arc, with very limited expression observed further up the third ventricle, in both genetic backgrounds.

In the SCN of P5 wildtype mice, very little expression of GFAP was observed (Figure 6.29). This is in contrast to the observation in adult wildtype mice that expression of GFAP is widespread in the SCN (Figure 6.16). Some expression of GFAP is observed around the third ventricle of this region, although the morphology of the projections appear more typical of the cells surrounding the third ventricle at the Arc, than those observed around the third ventricle near the SCN in adult mice. These observations suggest that the glial cells in this region are still developing, and that the cells may not have established themselves within the SCN at P5.

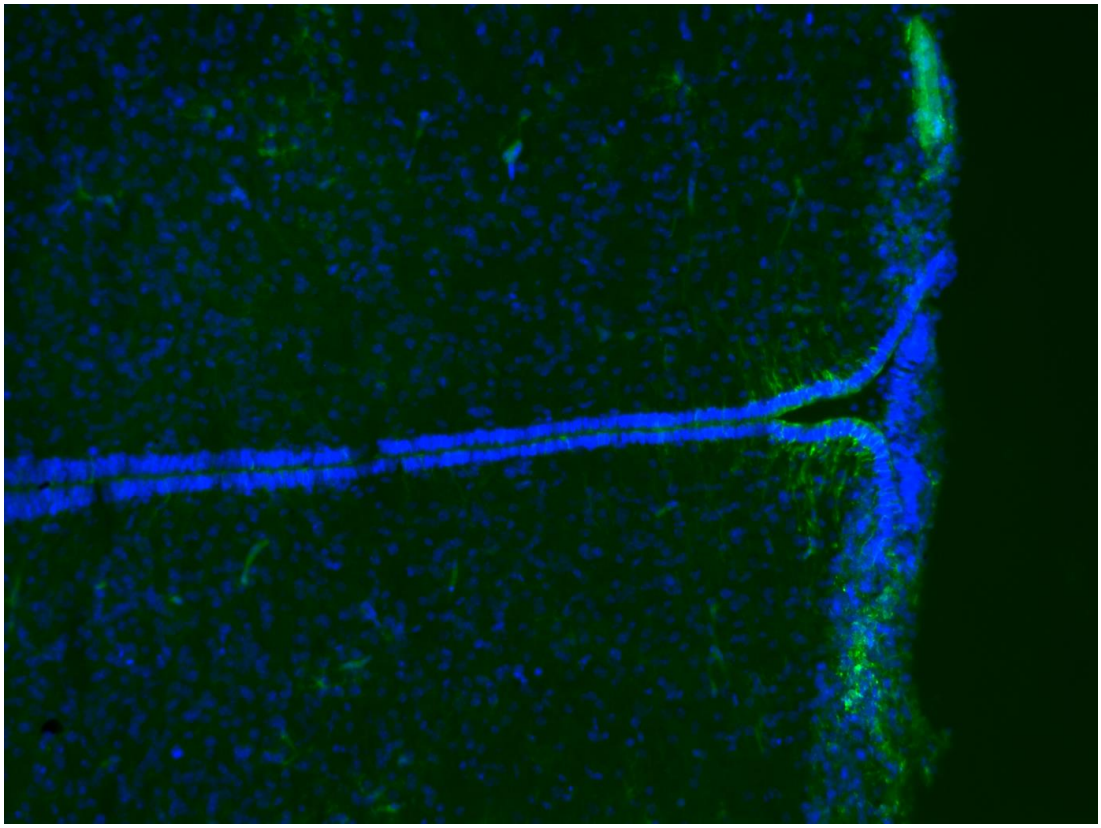

Figure 6.27: Immunofluorescence of the Arc of a P5 wildtype mouse

Immunofluorescence of the arcuate nucleus of the hypothalamus of a 14µm-thick coronal section of a wildtype mouse at P5 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

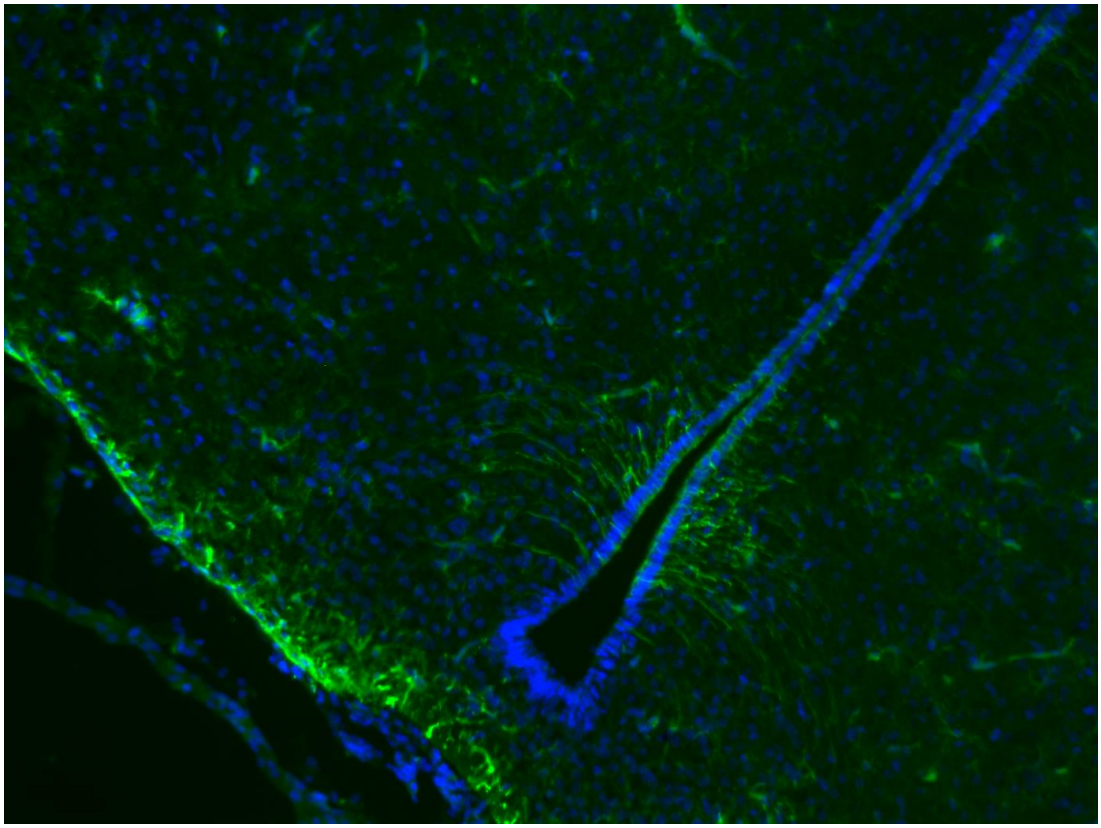

Figure 6.28: Immunofluorescence of the Arc of a P5 *Gnasxl*$^{m+/p-}$ mouse Immunofluorescence of the arcuate nucleus of the hypothalamus of a 14µm-thick coronal section of a *Gnasxl*$^{m+/p-}$ mouse at P5 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

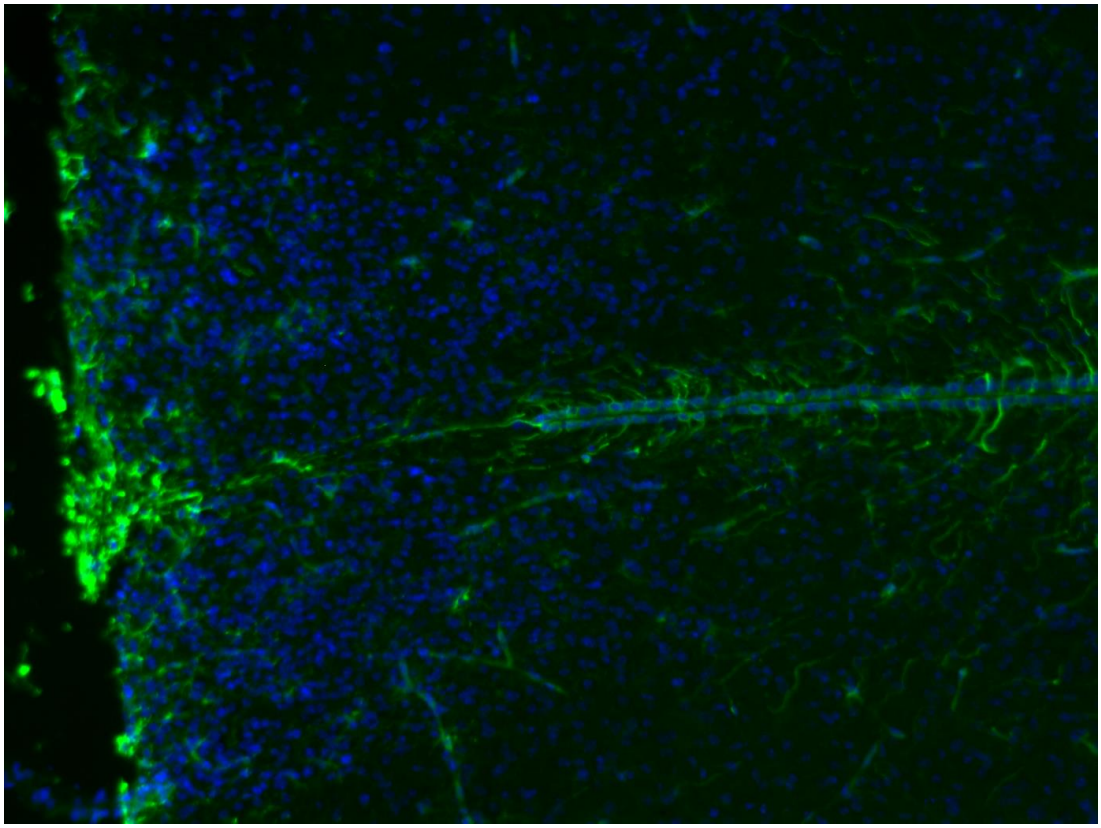

Figure 6.29: Immunofluorescence of the SCN of a P5 wildtype mouse

Immunofluorescence of the suprachiasmatic nucleus of the hypothalamus of a 14µm-thick coronal section of a wildtype mouse at P5 at 10X magnification. Sections were stained for GFAP (green) and DAPI (blue).

6.5 Discussion

Gfap was identified as being consistently downregulated two-fold in Gnasxl$^{m+/p-}$ mice in both the RNAseq and qPCR analyses. However, since mRNA levels do not always correlate with protein levels, we decided to analyse GFAP protein in the hypothalamus.

Western blot analyses were fairly inconclusive. Blots run on self-cast SDS-PAGE gels showed poorly defined bands compared to the NuPAGE SDS-PAGE Gel System (Life Technologies) (Figure 6.5, Figure 6.9). In particular for wildtype animals, a high degree of variability in the samples was observed (Figures 6.5, 6.6, 6.7, 6.8). Repeat experiments of the Western blots are advisable, due to the variability observed from blot to blot and the unexplained nature of three new wildtype samples, 123.2, 115.1, and 104.2 (Figure 6.9). Because these three samples had not been run on previous SDS-PAGE gels, it is not clear whether these samples truly have a dramatic loss of GFAP in hypothalamus lysates, or whether it was due to technical error.

The identification of both *H6pd* and *Hsd11b1* being upregulated 4-fold is particularly interesting due to their close functional connection in driving the activation of deoxycorticosterone into the active glucocorticoid corticosterone. Corticosterone is a glucocorticoid hormone that is produced in the adrenal glands and has a wide range of functions, including immune and stress responses, as well as energy regulation (Groeneweg et al., 2011). Corticosterone acts to inhibit the release of Corticotropin-releasing hormone (CRH), which is released from the PVN in response to stress. In particular, it is able to stimulate the action of POMC neurons, which act to reduce appetite. After release, CRH travels to the ME of the hypothalamus and then to the pituitary gland. Here, it stimulates corticotropic cells to produce adenocorticotropic hormone (ACTH). In turn, ACTH stimulates the production of corticosterone, creating a feedback loop (Bonfiglio et al., 2011). This forms the basis of the Hypothalamic-pituitary-adrenal (HPA) axis.

If corticosterone levels are elevated by the increased expression of *H6pd* and *Hsd11b1*, then this would act to suppress CRH. This could result in a reduced activation of POMC neurons by CRH, and therefore a reduction in its suppressive effects on appetite. Due to a reduction in CRH, ACTH may also be reduced, leading to less activation of corticosterone – allowing for only a moderate increase in corticosterone to be maintained. Gnasxl$^{m+/p-}$ mice show a modest increase in corticosterone levels, although this was not significant (Xie et al., 2006). It would

be interesting to see whether corticosterone levels are increased in a larger sample size of Gnasxl$^{m+/p-}$ mice, and to see whether CRH levels are reduced.

Importantly, corticosterone acts on neurons in the PVN by promoting the release of GABA and inhibiting the release of glutamate. The release of the inhibitory neurotransmitter GABA is mediated by the G protein βγ subunits, whereas preventing the release of the excitation neurotransmitter glutamate is regulated by Gαs (Di et al., 2005; Di et al., 2009). As an inhibitory neurotransmitter, GABA acts to suppress the action of POMC neurons in the PVN, promoting orexigenic effects.

Furthermore, Gαs is also known to couple to the CRH receptor (CRHR) to instigate downstream cAMP signalling to activate POMC expression (Bonfiglio et al., 2011; Hillhouse and Grammatopoulos, 2013). Whether XLαs has a role in this mechanism is not known. *In vitro*, XLαs is able to induce more sustained signalling due to its membrane anchoring (Liu et al., 2011). Therefore XLαs-coupled signalling could have wide implications in the feedback mechanism of the HPA axis. It would be interesting to know whether XLαs is able to facilitate the signalling cascade in response to corticosterone, or whether it is able to couple to CRHR like Gαs.

MC4R$^{-/-}$ mice are a genetic model of obesity. The lack of MC4R prevents the action of the inhibitory effects of POMC neurons on feeding and appetite – essentially shutting down the feedback mechanism, leading to overfeeding and weight gain. MC4R$^{-/-}$ mice display heightened levels of GFAP within the hypothalamus, particularly in the DMH and Arc, but also in the VMH and LH (Hsuchou et al., 2009; Buckman et al., 2013).

Conversely, *Gnasxl*$^{m+/p-}$ mice have an overactive SNS (Xie et al., 2006), which burns through the calories that the mice consume. Despite consuming more food, the mice remain small and lean and show similarities to chronic undernutrition (Plagge et al., 2004). It is therefore a parallel finding that immunofluorescence analysis suggests that *Gnasxl*$^{m+/p-}$ mice show a distinct loss of GFAP within the hypothalamus, particularly within the Arc, DMH, and SCN.

Using an inducible knockout driven by *Gfap*, the leptin receptor *LepR* was knocked out specifically in *Gfap*-expressing cells. Interestingly, the team identified that while the number of *Gfap*-expressing cells did not change, the number of cell projections, and the length of the projections were significantly decreased (Kim et al., 2014). Because the *Gfap*-expressing cells are unable to respond to leptin, food intake is significantly increased.

Collectively, it is hypothesised that leptin action on *Gfap*-expressing cells in the hypothalamus may instigate morphological changes in tanycytes and astrocytes. Given the extremely low leptin levels in *Gnasxl*$^{m+/p-}$ mice (Xie et al., 2006), it may be that this leptin-induced change does not occur to the same degree, and renders the hypothalamic glial cells smaller and with shorter projections.

In order to investigate whether there was a developmental problem causing a decrease in GFAP observed by immunofluorescence, early postnatal stages were analysed for GFAP levels in the hypothalamus. At P1 and P5 postnatal stages, no inhibition of gliogenesis was observed, with GFAP immunofluorescence within the hypothalamus appearing more or less the same for both wildtype and *Gnasxl*$^{m+/p-}$ mice. By P5, most differentiation of astrocytes has occurred (Sauvageot, 2002), which account for most of the glial cells in the hypothalamus. However, astrocytes are still migrating at this stage (Jacobsen and Miller, 2003). In addition, since gliogenesis continues for several weeks after birth, it would be interesting to identify at what stage *Gnasxl*$^{m+/p-}$ mice begin to show changes in GFAP levels and localisation within the hypothalamus. Unfortunately, due to time constraints of the project, it was not possible to study mice beyond P5. In addition, since immunofluorescence is not entirely quantitative, cell count experiments would help to identify whether the number of GFAP-positive cells is altered in *Gnasxl*$^{m+/p-}$ mice. Because of the morphology of astrocytes and tanycytes, it was difficult to distinguish cells from one another to provide this analysis. This can be achieved using the VectaStain ABC Kit protocol (Vector Laboratories). After binding of the anti-GFAP primary antibody to slide sections, a biotinylated secondary antibody is added, which has a high affinity for HRP-tagged Activin. After providing the VectaStain substrate, a brown stain appears identifying GFAP-positive cells. These can then be counted under light microscopy to identify any changes in GFAP-positive cell numbers in *Gnasxl*$^{m+/p-}$ mice.

The hypothalamus effects responses to various hormones. In female rats, a diurnal rhythm of *Gfap* in the SCN was identified that was enhanced upon treatment with estradiol benzoate, a synthetic form of the female hormone oestrogen (Fernandez-Galaz et al., 1999). The Arc responds to hormonal levels during the ovarian cycle, resulting in an increase in glial cell coverage, which acts to limit the action of the inhibitory GABAergic neurons (Garcia-Segura et al., 2008). For the purposes of our study, only male mice were used for RNAseq

and downstream analyses. Furthermore, littermates were sacrificed at similar times during the day, minimising the chance of expression changes.

During puberty, hormones such as testosterone and oestrogen act on glial cells within hypothalamus, in regions such as the SCN, Arc, and ME in order to regulate the production of gonadotropin-releasing hormone (GnRH) (Garcia-Segura et al., 2008). GnRH acts to stimulate the release of follicle-stimulating hormone and luteinizing hormone, which act to regulate testosterone production in males, and oestrogen production in females.

Glial cells have been implicated to have important roles within the SCN – the regulator of the circadian rhythm (Prosser et al., 1994). More recently it was identified that vesicle trafficking and calcium signalling within SCN astrocytes is crucial for glial cells to be able to modulate circadian rhythm activity (Ng et al., 2011).

Immunofluorescence analysis of the SCN suggested a distinct loss of glial cells within the SCN of *Gnasxl*$^{m+/p-}$ mice. No circadian rhythm abnormalities have been identified in *Gnasxl*$^{m+/p-}$ mice, although recent evidence points towards heart rate being significantly increased during the night (Nunn et al., 2013). However, it would be interesting to investigate whether the circadian rhythm of *Gnasxl*$^{m+/p-}$ mice is disrupted, particularly in adult mice by which time the differences in glial cell populations in the SCN are different.

Our results suggest that *Gnasxl*$^{m+/p-}$ mice show a distinct loss of GFAP within the Arc, and SCN. We find no evidence that gliogenesis is disrupted up to P5, whereby much of the astrocytic differentiation has already occurred. This suggests that the loss of GFAP in the hypothalami of *Gnasxl*$^{m+/p-}$ mice is unlikely to be a developmental disruption. Given the evidence linking the action of leptin to the morphology of *Gfap*-expressing cells in the hypothalamus (Kim et al., 2014), it is a more likely proposition that the loss of GFAP is due to the chronic undernutrition experienced by that *Gnasxl*$^{m+/p-}$ mice as a result of an overactive SNS.

Chapter 7: Discussion

In this study, we have used transcriptomic approaches to develop a better understanding of the complex biological phenomena of ageing and metabolism.

Using qPCR, it was possible to confirm the relative transcriptional level of genes in the NMR. Compared to mice, these genes were shown to be overexpressed. Since gene expression levels tend to be conserved between relatively distantly related species such as mice and humans (Zheng-Bradley et al., 2010), it is likely that such changes in expression between orthologous genes of closely related species may not be widespread. Therefore the genes identified could be relevant foci for future research into the interesting phenotypes of the NMR.

Further to qPCR analysis, the use of RNAseq analysis was widely employed in this study. Using SOLiD RNAseq combined with mismatch analysis, it was possible to identify *Cog3* as a conserved target of RNA editing in the rat. However, the finding that RNA editing does not decline in rat cerebral cortex is interesting, since this does occur in humans (Nicholas et al., 2010). This disparity between RNA editing decline in rats and humans highlights the important differences between humans, and their rodent model counterparts. In particular, there are differences in physiology, disease susceptibility, population demographics, and in their genetic responses to inflammation (Demetrius, 2005; Seok et al., 2013). Since rodent models are used to study ageing, it is important to understand how species differences in biological processes may affect the ageing process between species.

The development of new technologies, such as Digital Droplet (dd)PCR (Bio-Rad) (Hindson et al., 2011; Hindson et al., 2013) may provide new methods for RNA editing quantification. This technology splits the PCR reaction into around 20000 separate droplets. Some droplets will contain the target cDNA, and the proportion that do can be used for quantification. A recent study was able to use ddPCR to identify SNPs, demonstrating its practicality (Mukaide et al., 2014).

Using Illumina RNAseq, it was possible to identify gene expression changes within the hypothalami of *Gnasxl*$^{m+/p-}$ mice. Bioinformatic analysis highlighted a downregulation in *Gfap* RNA expression within the hypothalamus. After analysing GFAP protein expression by immunofluorescence, we were able to identify a loss of GFAP expression within hypothalamic regions, including the Arc and the SCN. No differences between wildtype and

Gnasxl$^{m+/p-}$ mice were observed during P1 and P5 stages. However, because these are still early stages of gliogenesis, it would be important to look at later stages of development (P10, P15, P20) in order to characterise when the difference in GFAP localisation arises. Our results suggest that the observed changes in GFAP expression are unlikely to be a disruption during early gliogenesis. The changes are more likely to be an adaptive response to chronic undernutrition that manifests itself in neonatal mice, a time in which the mice develop problems with suckling ability (Xie et al., 2006).

One of the main challenges with performing RNAseq on the hypothalamus was that as a complex region of the brain, there are many distinct but small nuclei that have diverse biological roles. Due to time constraints and financial costs, it was not possible to dissect out each nuclei and identify gene expression changes in the separate nuclei independently. However, this could be achieved using laser microdissection (Solga et al., 2015) in order to separate cell types from nuclei such as the Arc.

Standard RNAseq performed across a tissue has its own limitations. Across the hypothalamus, the various nuclei and cell types pose a challenge in identifying whether gene expression changes are as a result of decreased expression in cells, or a reduction in the number of cells. As technologies improve, single cell RNAseq is becoming more common. Techniques such as these have the benefit of being able to identify gene expression changes within a single cell. Combined with laser microdissection, single cell RNAseq may provide a basis for research into gene expression changes within distinct cell populations, such as the tanycytes

Recent advances have allowed the development of a technique that removes lipids from brain tissue, rendering it transparent. Using this technique, a new imaging method called CLARITY was developed whereby it is possible to perform immunofluorescence within the whole intact 3D mouse brain (Chung et al., 2013). The authors also performed immunofluorescence of *Gfap.* Rather than analysing sections of the mouse brain, as was performed in this project, in future it may be possible for expression patterns within tissues to be imaged and analysed across the whole tissue.

This project highlights the importance of new sequencing technologies in identifying gene expression changes. With the advent of new technologies, the amount of sequencing data being produced is increasing rapidly. Techniques such as RNAseq are not only generating more data and becoming more widespread, but costs are constantly falling. Furthermore, combining this with other advanced techniques has expanded the potential of RNAseq in

research. The continuing development of bioinformatic tools for gene expression analysis has great potential for identifying biologically relevant genetic changes. Consequently, transcriptomic analysis has a huge potential in directing future research, particularly in respect to complex research areas such as ageing and metabolism.

Appendices

6 month old samples

Target	8883F	8884F	8885F	8886R	8886F	8887R	8887F	8888F	STDEV	Mean
Gabra3	92.75	94.69			93.02		93.92	91.54	1.20	93.18
Cyfip2	88.32	88.44	90.93		88.39		89.23		1.10	89.06
Kcna1	22.31	25.76	27.80		37.21		31.91	22.27	5.84	27.87
Flna	21.15		25.91	28.18	24.71	26.11	23.77	21.32	2.58	24.45
Cog3	27.71	25.39			26.56		24.86	22.11	2.11	25.32
Blcap Y/C	46.59	44.33	39.73	38.69	41.70	35.09	36.57	41.97	3.85	40.58
Blcap Q/R	27.14	25.45	22.68	29.20	23.88	26.72	19.65	23.57	2.98	24.79
Blcap K/R	10.31	9.85	6.97		9.12		6.53	9.03	1.54	8.64

Appendix Table 1: Percentages of RNA editing identified in cerebral cortex from 6-month-old rats

F in the sample name indicates that it was sequenced using the forward primer, as given in Table 2.2. Conversely, R indicates it was sequenced using the reverse primer.

12 month old samples

Target	8910F	8911R	8911F	8912F	8913R	8913F	8914F	8915F	STDEV	Mean
Gabra3	94.35		91.59	91.72		93.30	93.09	92.24	1.06	92.71
Cyfip2	90.70		86.27	83.01		89.02	86.42	90.23	2.92	87.61
Kcna1			18.18	22.43		23.10	27.07	30.60	4.74	24.28
Flna	31.03	24.07	24.81	33.94	25.97	25.76	29.84	32.56	3.81	28.50
Cog3	21.60		24.23			24.92	26.67	20.00	2.67	23.48
Blcap Y/C	40.61	40.37	43.97	45.06	41.07	41.93	32.11	38.89	3.93	40.50
Blcap Q/R	22.43	30.71	24.20	27.51	31.94	22.80	20.18	20.92	4.45	25.09
Blcap K/R	6.95		9.65	10.63		8.90	8.80	7.12	1.43	8.67

Appendix Table 2: Percentages of RNA editing identified in cerebral cortex from 12-month-old rats

F in the sample name indicates that it was sequenced using the forward primer, as given in Table 2.2. Conversely, R indicates it was sequenced using the reverse primer.

28 month old samples

Target	8773R	8773F	8774F	8776F	8777R	8777F	8778F	8780F	STDEV	Mean
Gabra3		92.34	92.35	92.65		95.45			1.51	93.20
Cyfip2		86.04	88.31	87.58		89.30	87.27	89.57	1.33	88.01
Kcna1		20.44	27.68	23.05		35.88	34.50	36.92	7.04	29.74
Flna	26.02	24.90	23.26	26.04	29.41		27.41	28.25	2.08	26.47
Cog3		24.85		26.67		27.54	24.91	16.87	4.24	24.17
Blcap Y/C	41.08	43.33	41.85	43.81	37.65	39.96	41.38	35.11	2.91	40.52
Blcap Q/R	31.20	24.73	24.54	25.65	28.78	20.94	24.07	18.21	4.08	24.77
Blcap K/R		9.84	9.70	9.91		7.65	10.98	6.20	1.77	9.04

Appendix Table 3: Percentages of RNA editing identified in cerebral cortex from 28-month-old rats

F in the sample name indicates that it was sequenced using the forward primer, as given in Table 2.2. Conversely, R indicates it was sequenced using the reverse primer.

	One way ANOVA between ages	
Target	**p value**	**p value (Bonferroni correction)**
Gabra3	0.769	6.152
Cyfip2	0.492	3.936
Kcna1	0.346	2.768
Flna	0.052	0.416
Cog3	0.654	5.232
Blcap Y/C	0.999	7.992
Blcap Q/R	0.983	7.864
Blcap K/R	0.885	7.08

Appendix Table 4: ANOVA analysis of RNA editing percentages

One way ANOVA was performed, indicating that RNA editing was not significantly changed with age for any genes tested

gene	KO FPKM	WT FPKM	log2(fold_change)	q_value	Ensembl ID	Description
Gm10233	3.03865	0	-1.79769e+308	0.010555	ENSMUSG00000068165	predicted pseudogene 10233 [Source:MGI Symbol;Acc:MGI:3704447]
SNORA22	6131.66	444026	6.17822	0	ENSMUSG00000065304	Small nucleolar RNA SNORA22 [Source:RFAM;Acc:RF00414]
Snora30	44.1854	1449.09	5.03544	0.029654	ENSMUSG00000064581	Small nucleolar RNA SNORA30/SNORA37 family [Source:RFAM;Acc:RF00415]
SNORA79	111.094	3308.77	4.89644	3.18E-10	ENSMUSG00000088412	Small nucleolar RNA SNORA79 [Source:RFAM;Acc:RF00600]
SNORA35	266.42	5877.11	4.46333	8.60E-07	ENSMUSG00000084421	Small nucleolar RNA SNORA35 [Source:RFAM;Acc:RF00566]
Gm6166	7.82314	164.454	4.39379	0	ENSMUSG00000074280	predicted gene 6166 [Source:MGI Symbol;Acc:MGI:3645893]
SNORA70	110.628	1942.09	4.13382	0.018619	ENSMUSG00000087916	Small nucleolar RNA SNORA70 [Source:RFAM;Acc:RF00156]
Gm14934	4.22012	62.963	3.89915	1.47E-08	ENSMUSG00000082035	predicted gene 14934 [Source:MGI Symbol;Acc:MGI:3802129]
Gm2174	5.46429	65.9188	3.59259	2.16E-07	ENSMUSG00000058932	predicted gene 2174 [Source:MGI Symbol;Acc:MGI:3780344]
SNORA38	493.916	5402.24	3.45122	0	ENSMUSG00000065653	Small nucleolar RNA SNORA38 [Source:RFAM;Acc:RF00428]
Gm14305, Gm14306	0.501043	4.99975	3.31885	0.015848	#N/A	#N/A
SNORA48	46.6063	429.819	3.20513	0.03374	ENSMUSG00000089011	Small nucleolar RNA SNORA48 [Source:RFAM;Acc:RF00554]

Hist1h4c	2.84226	25.3616	3.15753	0.0032	ENSMUSG00000060678	histone cluster 1, H4c [Source:MGI Symbol;Acc:MGI:2448421]
Gm17383	6.41447	56.1118	3.1289	0	ENSMUSG00000058625	predicted gene, 17383 [Source:MGI Symbol;Acc:MGI:4937017]
SNORA70	18496.3	139409	2.91402	0	ENSMUSG00000087916	Small nucleolar RNA SNORA70 [Source:RFAM;Acc:RF00156]
Scarna6	70.2752	526.873	2.90637	4.78E-12	ENSMUSG00000087968	Small Cajal body specific RNA 6 [Source:RFAM;Acc:RF00478]
Gm4724	1.49152	9.36433	2.65039	0.012859	ENSMUSG00000078897	predicted gene 4724 [Source:MGI Symbol;Acc:MGI:3782904]
Gm10154	36.3201	172.951	2.25153	4.92E-13	ENSMUSG00000066116	predicted gene 10154 [Source:MGI Symbol;Acc:MGI:3642271]
Eif1	36.4411	166.612	2.19286	3.44E-07	ENSMUSG00000035530	eukaryotic translation initiation factor 1 [Source:MGI Symbol;Acc:MGI:105125]
Orc5	4.87409	21.7581	2.15834	0.036365	ENSMUSG00000029012	origin recognition complex, subunit 5 [Source:MGI Symbol;Acc:MGI:1347044]
Rpl31-ps9	33.8222	141.697	2.06677	8.47E-07	ENSMUSG00000094122	ribosomal protein L31, pseudogene 9 [Source:MGI Symbol;Acc:MGI:3704195]
H6pd	1.61313	6.49417	2.00928	0.033843	ENSMUSG00000028980	hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase) [Source:MGI Symbol;Acc:MGI:2140356]
Hsd11b1	6.44958	25.4623	1.98108	0.004499	ENSMUSG00000016194	hydroxysteroid 11-beta dehydrogenase 1 [Source:MGI Symbol;Acc:MGI:103562]

Gm9844	5.46518	21.3509	1.96596	0.005045	ENSMUSG00000050347	predicted pseudogene 9844 [Source:MGI Symbol;Acc:MGI:3704288]
Rps11	69.737	270.368	1.95493	0	ENSMUSG00000003429	ribosomal protein S11 [Source:MGI Symbol;Acc:MGI:1351329]
Snapc5	5.21811	20.1511	1.94926	5.15E-07	ENSMUSG00000032398	small nuclear RNA activating complex, polypeptide 5 [Source:MGI Symbol;Acc:MGI:1914282]
Rpl22l1	20.9724	79.7795	1.92753	2.87E-08	ENSMUSG00000039221	ribosomal protein L22 like 1 [Source:MGI Symbol;Acc:MGI:1915278]
Gm9892	0.93136	3.43611	1.88337	0.007047	ENSMUSG00000052825	predicted gene 9892 [Source:MGI Symbol;Acc:MGI:3701610]
Phb2	57.9534	211.71	1.86913	9.67E-06	ENSMUSG00000004264	prohibitin 2 [Source:MGI Symbol;Acc:MGI:102520]
Ambra1	25.9227	93.18	1.8458	0.043294	ENSMUSG00000040506	autophagy/beclin 1 regulator 1 [Source:MGI Symbol;Acc:MGI:2443564]
Gm13706	30.8538	106.719	1.7903	0.010687	ENSMUSG00000085838	predicted gene 13706 [Source:MGI Symbol;Acc:MGI:3702041]
Gm14586	5.24509	17.6781	1.75292	0.026534	ENSMUSG00000083621	predicted gene 14586 [Source:MGI Symbol;Acc:MGI:3705507]
Smek2	16.5215	54.0715	1.71052	3.80E-06	ENSMUSG00000020463	SMEK homolog 2, suppressor of mek1 (Dictyostelium) [Source:MGI Symbol;Acc:MGI:2144474]
Pdap1	28.1594	91.5695	1.70125	5.44E-12	ENSMUSG00000029623	PDGFA associated protein 1 [Source:MGI Symbol;Acc:MGI:2448536]

Gm8430	10.8643	34.8721	1.68248	0.003336	ENSMUSG00000055093	predicted pseudogene 8430 [Source:MGI Symbol;Acc:MGI:3645406]
Gm561	12.3768	39.2545	1.66522	0.001446	ENSMUSG00000074754	predicted gene 561 [Source:MGI Symbol;Acc:MGI:2685407]
Tmem101	2.30856	7.31026	1.66293	0.013841	ENSMUSG00000020921	transmembrane protein 101 [Source:MGI Symbol;Acc:MGI:1923797]
Gm13072	10.8166	34.029	1.65352	0.000127	ENSMUSG00000070572	predicted gene 13072 [Source:MGI Symbol;Acc:MGI:3651720]
Rps29	210.337	648.823	1.62512	0	ENSMUSG00000034892	ribosomal protein S29 [Source:MGI Symbol;Acc:MGI:107681]
Gm2026	1.58502	4.84516	1.61204	0.043396	ENSMUSG00000078886	predicted gene 2026 [Source:MGI Symbol;Acc:MGI:3780195]
Rpl39	91.4318	279.247	1.61077	0	ENSMUSG00000079641	ribosomal protein L39 [Source:MGI Symbol;Acc:MGI:1914498]
Ndufb3	6.15747	18.5638	1.59208	0.001907	ENSMUSG00000026032	NADH dehydrogenase (ubiquinone) 1 beta subcomplex 3 [Source:MGI Symbol;Acc:MGI:1913745]
2610001J05Rik	8.50419	25.5924	1.58947	1.91E-05	ENSMUSG00000052419	RIKEN cDNA 2610001J05 gene [Source:MGI Symbol;Acc:MGI:1913770]
AY036118	6.65412	19.7018	1.566	0.005605	ENSMUSG00000045999	cDNA sequence AY036118 [Source:MGI Symbol;Acc:MGI:2158419]
Hist2h4	1.70069	4.99366	1.55398	0.03374	ENSMUSG00000091405	histone cluster 2, H4 [Source:MGI Symbol;Acc:MGI:2140113]
Rimklb	9.37839	27.2163	1.53706	0.015824	ENSMUSG00000040649	ribosomal modification protein rimK-like family member B [Source:MGI Symbol;Acc:MGI:1918325]

Gm4294	7.15543	20.6122	1.52639	0.021302	ENSMUSG00000078377	predicted gene 4294 [Source:MGI Symbol;Acc:MGI:3782472]
SNORA8	295.273	822.076	1.47722	0.015147	ENSMUSG00000065226	Small nucleolar RNA SNORA8 [Source:RFAM;Acc:RF00393]
Drap1	19.1003	52.3954	1.45584	0.001026	ENSMUSG00000024914	Dr1 associated protein 1 (negative cofactor 2 alpha) [Source:MGI Symbol;Acc:MGI:1913806]
Cstb	6.74401	18.2706	1.43785	0.013774	ENSMUSG00000005054	cystatin B [Source:MGI Symbol;Acc:MGI:109514]
Gm6142	123.045	331.793	1.4311	1.59E-06	ENSMUSG00000084304	predicted pseudogene 6142 [Source:MGI Symbol;Acc:MGI:3644670]
Rpl21-ps15	36.0916	96.8072	1.42345	0.004073	ENSMUSG00000083596	ribosomal protein L21, pseudogene 15 [Source:MGI Symbol;Acc:MGI:3705426]
Rpl35a	118.761	316.151	1.41256	1.28E-06	ENSMUSG00000060636	ribosomal protein L35A [Source:MGI Symbol;Acc:MGI:1928894]
Gabarap	63.1082	167.447	1.4078	7.26E-07	ENSMUSG00000018567	gamma-aminobutyric acid receptor associated protein [Source:MGI Symbol;Acc:MGI:1861742]
Scarna13	199.094	513.302	1.36636	1.30E-08	ENSMUSG00000088789	small Cajal body-specific RNA 1 [Source:MGI Symbol;Acc:MGI:4360027]
Rps21	267.96	680.529	1.34464	1.45E-11	ENSMUSG00000039001	ribosomal protein S21 [Source:MGI Symbol;Acc:MGI:1913731]
1810027O10Rik	27.8734	70.6814	1.34244	0.035806	ENSMUSG00000070394	RIKEN cDNA 1810027O10 gene [Source:MGI Symbol;Acc:MGI:1916436]

Rpl39-ps	86.5495	219.384	1.34186	3.37E-07	ENSMUSG00000036305	ribosomal protein L39, pseudogene [Source:MGI Symbol;Acc:MGI:3705781]
mt-Nd3	1932.19	4841.1	1.3251	0	ENSMUSG00000064360	mitochondrially encoded NADH dehydrogenase 3 [Source:MGI Symbol;Acc:MGI:102499]
2010107E04Rik	35.4377	88.6033	1.32208	0.00211	ENSMUSG00000021290	RIKEN cDNA 2010107E04 gene [Source:MGI Symbol;Acc:MGI:1917507]
Mt2	13.6123	33.9086	1.31674	0.01513	ENSMUSG00000031762	metallothionein 2 [Source:MGI Symbol;Acc:MGI:97172]
Hbb-b1	20.9713	52.2282	1.31641	0.001917	ENSMUSG00000052305	hemoglobin, beta adult major chain [Source:MGI Symbol;Acc:MGI:96021]
Rpl31-ps8	70.577	174.753	1.30804	8.30E-07	ENSMUSG00000067870	ribosomal protein L31, pseudogene 8 [Source:MGI Symbol;Acc:MGI:3647726]
Ccdc132	22.5027	55.641	1.30605	0.027976	ENSMUSG00000001376	coiled-coil domain containing 132 [Source:MGI Symbol;Acc:MGI:1920538]
Letm2	4.64771	11.4583	1.3018	0.0003	ENSMUSG00000037363	leucine zipper-EF-hand containing transmembrane protein 2 [Source:MGI Symbol;Acc:MGI:2444979]
Gm6457	10.993	26.8361	1.2876	0.000633	ENSMUSG00000053740	predicted pseudogene 6457 [Source:MGI Symbol;Acc:MGI:3648519]
Usmg5	111.628	268.848	1.2681	1.49E-05	ENSMUSG00000071528	upregulated during skeletal muscle growth 5 [Source:MGI Symbol;Acc:MGI:1891435]

Vps29	44.4335	106.632	1.26292	0.013862	ENSMUSG00000029462	vacuolar protein sorting 29 (S. pombe) [Source:MGI Symbol;Acc:MGI:1928344]
Rpl31	125.854	300.813	1.25711	7.36E-11	ENSMUSG00000073702	ribosomal protein L31 [Source:MGI Symbol;Acc:MGI:2149632]
Gm6356	2.90718	6.94572	1.2565	0.023087	ENSMUSG00000091400	predicted gene 6356 [Source:MGI Symbol;Acc:MGI:3646390]
Tmsb4x	224.954	534.658	1.24899	2.04E-11	ENSMUSG00000049775	thymosin, beta 4, X chromosome [Source:MGI Symbol;Acc:MGI:99510]
Ube2w	53.5845	125.116	1.22338	0.002575	ENSMUSG00000025939	ubiquitin-conjugating enzyme E2W (putative) [Source:MGI Symbol;Acc:MGI:1914049]
Mrps18c	22.7765	52.6535	1.20898	0.001599	ENSMUSG00000016833	mitochondrial ribosomal protein S18C [Source:MGI Symbol;Acc:MGI:1915985]
Srsf3	53.5626	123.074	1.20023	6.73E-06	ENSMUSG00000071172	serine/arginine-rich splicing factor 3 [Source:MGI Symbol;Acc:MGI:98285]
Rps5	58.262	133.069	1.19154	8.22E-07	ENSMUSG00000012848	ribosomal protein S5 [Source:MGI Symbol;Acc:MGI:1097682]
Tbca	13.959	31.7489	1.18551	0.012646	ENSMUSG00000042043	tubulin cofactor A [Source:MGI Symbol;Acc:MGI:107549]
Cox7a2	37.9082	85.9804	1.1815	1.51E-05	ENSMUSG00000032330	cytochrome c oxidase, subunit VIIa 2 [Source:MGI Symbol;Acc:MGI:1316715]
Snrpf	14.9816	33.9021	1.17818	5.39E-05	ENSMUSG00000020018	small nuclear ribonucleoprotein polypeptide F [Source:MGI Symbol;Acc:MGI:1917128]
Cox7c	41.9201	94.4257	1.17154	0.00973	ENSMUSG00000017778	cytochrome c oxidase, subunit VIIc [Source:MGI Symbol;Acc:MGI:103226]

Ndufa3	163.327	361.38	1.14575	0.000159	ENSMUSG00000035674	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3 [Source:MGI Symbol;Acc:MGI:1913341]
Gm14303	56.5841	122.917	1.11922	0.033843	ENSMUSG00000081344	predicted gene 14303 [Source:MGI Symbol;Acc:MGI:3651379]
Hbb-b2	17.0414	36.6542	1.10494	0.01403	ENSMUSG00000073940	hemoglobin, beta adult minor chain [Source:MGI Symbol;Acc:MGI:96022]
Hnrnpa2b1	48.8077	104.939	1.10437	7.31E-06	ENSMUSG00000004980	heterogeneous nuclear ribonucleoprotein A2/B1 [Source:MGI Symbol;Acc:MGI:104819]
C430048L16Rik	5.0558	10.8668	1.10392	0.009706	ENSMUSG00000052137	RIKEN cDNA C430048L16 gene [Source:MGI Symbol;Acc:MGI:1924854]
Tceal6	22.4886	48.2993	1.10281	2.03E-07	ENSMUSG00000031409	transcription elongation factor A (SII)-like 6 [Source:MGI Symbol;Acc:MGI:1923939]
Uqcrb	250.207	535.237	1.09706	1.34E-07	ENSMUSG00000021520	ubiquinol-cytochrome c reductase binding protein [Source:MGI Symbol;Acc:MGI:1914780]
Rasa2	6.90981	14.7619	1.09516	0.009276	ENSMUSG00000032413	RAS p21 protein activator 2 [Source:MGI Symbol;Acc:MGI:2149960]
Kpnb1	31.6232	66.7009	1.07672	0.025467	ENSMUSG00000001440	karyopherin (importin) beta 1 [Source:MGI Symbol;Acc:MGI:107532]
Pfdn4	36.5132	76.399	1.06513	0.044235	ENSMUSG00000052033	prefoldin 4 [Source:MGI Symbol;Acc:MGI:1923512]
Rpl34-ps1	372.358	778.99	1.06491	7.14E-09	ENSMUSG00000068396	ribosomal protein L34, pseudogene 1 [Source:MGI Symbol;Acc:MGI:3704270]

Anapc13	36.4614	75.8889	1.05752	0.015888	ENSMUSG00000035048	anaphase promoting complex subunit 13 [Source:MGI Symbol;Acc:MGI:1916260]
Tmem181a	4.70879	9.751	1.05019	0.034494	ENSMUSG00000038141	transmembrane protein 181A [Source:MGI Symbol;Acc:MGI:1924356]
Gm7887	11.5446	23.7491	1.04065	0.024464	ENSMUSG00000044211	predicted gene 7887 [Source:MGI Symbol;Acc:MGI:3648251]
Capza2	118.33	243.243	1.03958	0.003418	ENSMUSG00000015733	capping protein (actin filament) muscle Z-line, alpha 2 [Source:MGI Symbol;Acc:MGI:106222]
Shfm1	103.341	212.416	1.03948	8.89E-06	ENSMUSG00000042541	split hand/foot malformation (ectrodactyly) type 1 [Source:MGI Symbol;Acc:MGI:109238]
Tm9sf4	21.1106	43.2414	1.03445	0.043728	ENSMUSG00000068040	transmembrane 9 superfamily protein member 4 [Source:MGI Symbol;Acc:MGI:2139220]
Ptp4a2	72.489	147.56	1.02546	0.012245	ENSMUSG00000028788	protein tyrosine phosphatase 4a2 [Source:MGI Symbol;Acc:MGI:1277117]
Arl6ip1	45.7296	92.9429	1.02321	1.20E-07	ENSMUSG00000030654	ADP-ribosylation factor-like 6 interacting protein 1 [Source:MGI Symbol;Acc:MGI:1858943]
Ddt, Gm20441, Gstt3	40.9111	82.7838	1.01686	0.002033	#N/A	#N/A
Dym	3.47473	7.03073	1.01678	0.039478	ENSMUSG00000035765	dymeclin [Source:MGI Symbol;Acc:MGI:1918480]
Prdx1	94.178	189.872	1.01157	0.001604	ENSMUSG00000028691	peroxiredoxin 1 [Source:MGI Symbol;Acc:MGI:99523]

Ndufs6	40.9114	82.2007	1.00665	0.011947	ENSMUSG00000021606	NADH dehydrogenase (ubiquinone) Fe-S protein 6 [Source:MGI Symbol;Acc:MGI:107932]
Ctxn2	9.00905	18.0793	1.00489	0.031113	ENSMUSG00000074872	cortexin 2 [Source:MGI Symbol;Acc:MGI:2139444]
Gm10335	48.043	96.1099	1.00036	0.001277	ENSMUSG00000071532	predicted gene 10335 [Source:MGI Symbol;Acc:MGI:3641693]
Zik1	1.8657	3.71838	0.994957	0.049002	ENSMUSG00000030393	zinc finger protein interacting with K protein 1 [Source:MGI Symbol;Acc:MGI:108070]
Gm10132	46.153	91.0523	0.98027	0.007039	ENSMUSG00000063556	predicted gene 10132 [Source:MGI Symbol;Acc:MGI:3704449]
Tmx2	93.6766	184.634	0.978903	0.011491	ENSMUSG00000050043	thioredoxin-related transmembrane protein 2 [Source:MGI Symbol;Acc:MGI:1914208]
Fdx1l	18.0773	34.9096	0.949449	0.019322	ENSMUSG00000079677	ferredoxin 1-like [Source:MGI Symbol;Acc:MGI:1915415]
Rpl5	210.84	403.644	0.936937	0.001907	ENSMUSG00000058558	ribosomal protein L5 [Source:MGI Symbol;Acc:MGI:102854]
Gm10119	45.577	87.2075	0.936147	6.44E-05	ENSMUSG00000062611	predicted gene 10119 [Source:MGI Symbol;Acc:MGI:3642853]
Calr	109.166	206.546	0.919936	0.002605	ENSMUSG00000003814	calreticulin [Source:MGI Symbol;Acc:MGI:88252]
Rps27	277.021	524.081	0.919795	0.00109	ENSMUSG00000090733	ribosomal protein S27 [Source:MGI Symbol;Acc:MGI:1888676]
Rpl15-ps3	57.7383	108.658	0.912195	0.000862	ENSMUSG00000061167	ribosomal protein L15, pseudogene 3 [Source:MGI Symbol;Acc:MGI:3782952]
Pfdn2	100.092	187.93	0.908865	0.007047	ENSMUSG00000006412	prefoldin 2 [Source:MGI Symbol;Acc:MGI:1276111]

Gm13841	60.335	112.576	0.899827	0.000573	ENSMUSG00000083833	predicted gene 13841 [Source:MGI Symbol;Acc:MGI:3650890]
Tmsb10	180.094	330.458	0.875711	0.000116	ENSMUSG00000079523	thymosin, beta 10 [Source:MGI Symbol;Acc:MGI:109146]
Cbln4	7.18924	13.1747	0.873858	0.011491	ENSMUSG00000067578	cerebellin 4 precursor protein [Source:MGI Symbol;Acc:MGI:2154433]
Plp1	154.56	281.253	0.863705	0.001907	ENSMUSG00000031425	proteolipid protein (myelin) 1 [Source:MGI Symbol;Acc:MGI:97623]
RNase_MRP	472.047	857.457	0.861135	0.000647	ENSMUSG00000088375	RNase MRP [Source:RFAM;Acc:RF00030]
Bex1	61.5529	111.416	0.856064	8.07E-05	ENSMUSG00000050071	brain expressed gene 1 [Source:MGI Symbol;Acc:MGI:1328321]
AC241392.1	6.13797	11.0855	0.852837	0.007047	ENSMUSG00000094614	0
Cetn2	29.9507	53.568	0.83878	0.026046	ENSMUSG00000031347	centrin 2 [Source:MGI Symbol;Acc:MGI:1347085]
Prkar1b	86.8814	153.879	0.824674	0.033843	ENSMUSG00000025855	protein kinase, cAMP dependent regulatory, type I beta [Source:MGI Symbol;Acc:MGI:97759]
Rpl34	108.379	190.659	0.814917	0.004624	ENSMUSG00000062006	ribosomal protein L34 [Source:MGI Symbol;Acc:MGI:1915686]
Sec62	42.6543	74.3601	0.801839	2.05E-05	ENSMUSG00000027706	SEC62 homolog (S. cerevisiae) [Source:MGI Symbol;Acc:MGI:1916526]
Rpl11	160.453	278.892	0.797557	0.014478	ENSMUSG00000059291	ribosomal protein L11 [Source:MGI Symbol;Acc:MGI:1914275]
Rpl29	60.6666	105.376	0.796577	0.002923	ENSMUSG00000048758	ribosomal protein L29 [Source:MGI Symbol;Acc:MGI:99687]

Tuba1a	74.4617	128.929	0.792004	0.003418	ENSMUSG00000072235	tubulin, alpha 1A [Source:MGI Symbol;Acc:MGI:98869]
2410004A20Rik, Gm20509, Rsl24d1	31.4957	54.4404	0.789523	0.034225	#N/A	#N/A
Eef1a1	726.75	1253.78	0.786756	0.008062	ENSMUSG00000037742	eukaryotic translation elongation factor 1 alpha 1 [Source:MGI Symbol;Acc:MGI:1096881]
Ube2d2	7.28257	12.5072	0.78024	0.048131	ENSMUSG00000091896	ubiquitin-conjugating enzyme E2D 2 [Source:MGI Symbol;Acc:MGI:1930715]
Rps9	116.409	199.609	0.777976	0.00352	ENSMUSG00000006333	ribosomal protein S9 [Source:MGI Symbol;Acc:MGI:1924096]
Gm10163	218.514	373.094	0.771809	0.002727	ENSMUSG00000066443	predicted pseudogene 10163 [Source:MGI Symbol;Acc:MGI:3704341]
Rps23	400.956	683.473	0.769439	7.71E-05	ENSMUSG00000049517	ribosomal protein S23 [Source:MGI Symbol;Acc:MGI:1913725]
Txn1	38.2081	64.8734	0.76375	0.009075	ENSMUSG00000028367	thioredoxin 1 [Source:MGI Symbol;Acc:MGI:98874]
Rpl10	141.596	236.818	0.742002	0.003238	ENSMUSG00000008682	ribosomal protein L10 [Source:MGI Symbol;Acc:MGI:105943]
Vbp1	56.5878	94.4358	0.738844	0.002151	ENSMUSG00000031197	von Hippel-Lindau binding protein 1 [Source:MGI Symbol;Acc:MGI:1333804]
Ppp2r2a	7.08438	11.6517	0.717824	0.018619	ENSMUSG00000022052	protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), alpha isoform [Source:MGI Symbol;Acc:MGI:1919228]
Gm12346	30.3403	49.7505	0.713473	0.002634	ENSMUSG00000083899	predicted gene 12346 [Source:MGI Symbol;Acc:MGI:3649810]

Rps17	262.317	429.926	0.712774	0.002836	ENSMUSG00000061787	ribosomal protein S17 [Source:MGI Symbol;Acc:MGI:1309526]
Morc3	12.0822	19.7821	0.711315	0.004406	ENSMUSG00000039456	microrchidia 3 [Source:MGI Symbol;Acc:MGI:2136841]
Vcpip1	3.83686	6.27119	0.708814	0.015848	ENSMUSG00000045210	valosin containing protein (p97)/p47 complex interacting protein 1 [Source:MGI Symbol;Acc:MGI:1917925]
Gm15500	112.217	182.085	0.698324	0.000778	ENSMUSG00000086583	predicted pseudogene 15500 [Source:MGI Symbol;Acc:MGI:3782947]
Ngfrap1	85.7991	139.218	0.698311	0.014301	ENSMUSG00000046432	nerve growth factor receptor (TNFRSF16) associated protein 1 [Source:MGI Symbol;Acc:MGI:1338016]
Ndufa6	78.6703	127.388	0.695335	0.010156	ENSMUSG00000022450	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6 (B14) [Source:MGI Symbol;Acc:MGI:1914380]
Hnrnpa1	74.3747	120.23	0.692907	0.005728	ENSMUSG00000046434	heterogeneous nuclear ribonucleoprotein A1 [Source:MGI Symbol;Acc:MGI:104820]
Naa15	7.95862	12.7386	0.678611	0.010791	ENSMUSG00000063273	N(alpha)-acetyltransferase 15, NatA auxiliary subunit [Source:MGI Symbol;Acc:MGI:1922088]
Cox5b	386.243	609.43	0.657952	0.002129	ENSMUSG00000061518	cytochrome c oxidase, subunit Vb [Source:MGI Symbol;Acc:MGI:88475]
Cox6c	584.109	918.613	0.65322	0.008607	ENSMUSG00000014313	cytochrome c oxidase, subunit VIc [Source:MGI Symbol;Acc:MGI:104614]

Ilf2	24.8897	39.0065	0.648164	0.018619	ENSMUSG00000001016	interleukin enhancer binding factor 2 [Source:MGI Symbol;Acc:MGI:1915031]
Asns	31.781	49.6131	0.642558	0.037589	ENSMUSG00000029752	asparagine synthetase [Source:MGI Symbol;Acc:MGI:1350929]
Eif3g	33.4211	52.1072	0.640723	0.017538	ENSMUSG00000070319	eukaryotic translation initiation factor 3, subunit G [Source:MGI Symbol;Acc:MGI:1858258]
Hnrnpab	79.0291	122.839	0.636314	0.013952	ENSMUSG00000020358	heterogeneous nuclear ribonucleoprotein A/B [Source:MGI Symbol;Acc:MGI:1330294]
Fam20b	26.3311	40.535	0.622398	0.015848	ENSMUSG00000033557	family with sequence similarity 20, member B [Source:MGI Symbol;Acc:MGI:2443990]
Gmfb	32.5982	50.0192	0.617689	0.007331	ENSMUSG00000062014	glia maturation factor, beta [Source:MGI Symbol;Acc:MGI:1927133]
Nop58	14.129	21.672	0.617174	0.043294	ENSMUSG00000026020	NOP58 ribonucleoprotein homolog (yeast) [Source:MGI Symbol;Acc:MGI:1933184]
Otub1	37.3088	56.8273	0.607065	0.038156	ENSMUSG00000024767	OTU domain, ubiquitin aldehyde binding 1 [Source:MGI Symbol;Acc:MGI:2147616]
Cwc15	45.4422	68.3745	0.589425	0.015848	ENSMUSG00000004096	CWC15 homolog (S. cerevisiae) [Source:MGI Symbol;Acc:MGI:1913320]
Cacna2d1	21.877	32.7972	0.58416	0.034494	ENSMUSG00000040118	calcium channel, voltage-dependent, alpha2/delta subunit 1 [Source:MGI Symbol;Acc:MGI:88295]

Tuba1b	86.8187	129.973	0.582136	0.019159	ENSMUSG00000023004	tubulin, alpha 1B [Source:MGI Symbol;Acc:MGI:107804]
Purb	21.0907	31.5531	0.581173	0.010233	ENSMUSG00000094483	purine rich element binding protein B [Source:MGI Symbol;Acc:MGI:1338779]
Sult4a1	42.0334	62.1917	0.565185	0.024626	ENSMUSG00000018865	sulfotransferase family 4A, member 1 [Source:MGI Symbol;Acc:MGI:1888971]
Ptma	45.7414	67.674	0.565103	0.028202	ENSMUSG00000026238	prothymosin alpha [Source:MGI Symbol;Acc:MGI:97803]
Rpl37	313.757	462.925	0.561129	0.018826	ENSMUSG00000041841	ribosomal protein L37 [Source:MGI Symbol;Acc:MGI:1914531]
Rbm6	14.1884	20.9257	0.560563	0.031079	ENSMUSG00000032582	RNA binding motif protein 6 [Source:MGI Symbol;Acc:MGI:1338037]
Atxn10	54.6095	78.5801	0.525013	0.034402	ENSMUSG00000016541	ataxin 10 [Source:MGI Symbol;Acc:MGI:1859293]
Nbea	12.8397	18.1241	0.4973	0.049903	ENSMUSG00000027799	neurobeachin [Source:MGI Symbol;Acc:MGI:1347075]
Aplp1	116.634	80.105	-0.54202	0.019919	ENSMUSG00000006651	amyloid beta (A4) precursor-like protein 1 [Source:MGI Symbol;Acc:MGI:88046]
Sst	251.568	172.569	-0.54378	0.049002	ENSMUSG00000004366	somatostatin [Source:MGI Symbol;Acc:MGI:98326]
Ptn	63.7929	43.4972	-0.55248	0.033843	ENSMUSG00000029838	pleiotrophin [Source:MGI Symbol;Acc:MGI:97804]
Epb4.1l3	23.4734	15.9276	-0.5595	0.042457	ENSMUSG00000024044	erythrocyte protein band 4.1-like 3 [Source:MGI Symbol;Acc:MGI:103008]

Rasgef1a	29.3328	19.7926	-0.56755	0.043261	ENSMUSG00000030134	RasGEF domain family, member 1A [Source:MGI Symbol;Acc:MGI:1917977]
Kcnk1	41.902	28.0906	-0.57694	0.049002	ENSMUSG00000033998	potassium channel, subfamily K, member 1 [Source:MGI Symbol;Acc:MGI:109322]
Ttc7b	25.4221	16.9756	-0.58262	0.049002	ENSMUSG00000033530	tetratricopeptide repeat domain 7B [Source:MGI Symbol;Acc:MGI:2144724]
Srp54a	48.1227	32.0755	-0.58525	0.030198	ENSMUSG00000021020	signal recognition particle 54A [Source:MGI Symbol;Acc:MGI:1346087]
mt-Nd2	1021.37	678.702	-0.58966	0.015205	ENSMUSG00000064345	mitochondrially encoded NADH dehydrogenase 2 [Source:MGI Symbol;Acc:MGI:102500]
Laptm4a	201.714	133.956	-0.59056	0.009268	ENSMUSG00000020585	lysosomal-associated protein transmembrane 4A [Source:MGI Symbol;Acc:MGI:108017]
Myl12b	98.5096	64.8239	-0.60374	0.049002	ENSMUSG00000034868	myosin, light chain 12B, regulatory [Source:MGI Symbol;Acc:MGI:107494]
Ckmt1	185.074	121.682	-0.60498	0.033843	ENSMUSG00000000308	creatine kinase, mitochondrial 1, ubiquitous [Source:MGI Symbol;Acc:MGI:99441]
Tmem66	168.961	110.912	-0.60727	0.007041	ENSMUSG00000031532	transmembrane protein 66 [Source:MGI Symbol;Acc:MGI:1915137]
Atp1b1	437.145	286.516	-0.6095	0.010233	ENSMUSG00000026576	ATPase, Na+/K+ transporting, beta 1 polypeptide [Source:MGI Symbol;Acc:MGI:88108]

Wbp4	18.4138	11.9757	-0.62068	0.047261	ENSMUSG00000022023	WW domain binding protein 4 [Source:MGI Symbol;Acc:MGI:109568]
Dctn2	44.5704	28.9034	-0.62485	0.022352	ENSMUSG00000025410	dynactin 2 [Source:MGI Symbol;Acc:MGI:107733]
Cfl1	222.766	143.674	-0.63273	0.004499	ENSMUSG00000056201	cofilin 1, non-muscle [Source:MGI Symbol;Acc:MGI:101757]
Gm15772	1181.45	760.353	-0.63582	0.0032	ENSMUSG00000062353	predicted gene 15772 [Source:MGI Symbol;Acc:MGI:3805541]
Rpl18a	85.5791	55.0297	-0.63705	0.018619	ENSMUSG00000045128	ribosomal protein L18A [Source:MGI Symbol;Acc:MGI:1924058]
Pnmal1	48.6174	31.251	-0.63757	0.01593	ENSMUSG00000041141	PNMA-like 1 [Source:MGI Symbol;Acc:MGI:1918941]
6330403A02Rik	8.98449	5.72008	-0.6514	0.038439	ENSMUSG00000053963	RIKEN cDNA 6330403A02 gene [Source:MGI Symbol;Acc:MGI:2138735]
Fads2	62.3819	39.683	-0.65261	0.020392	ENSMUSG00000024665	fatty acid desaturase 2 [Source:MGI Symbol;Acc:MGI:1930079]
Snx19	9.15313	5.81575	-0.6543	0.044932	ENSMUSG00000031993	sorting nexin 19 [Source:MGI Symbol;Acc:MGI:1921581]
Csnk1g2	29.7967	18.9259	-0.65479	0.034402	ENSMUSG00000003345	casein kinase 1, gamma 2 [Source:MGI Symbol;Acc:MGI:1920014]
Pja1	88.3012	55.9293	-0.65883	0.026114	ENSMUSG00000034403	praja1, RING-H2 motif containing [Source:MGI Symbol;Acc:MGI:1101765]
Kifc2	35.691	22.5938	-0.65963	0.008894	ENSMUSG00000004187	kinesin family member C2 [Source:MGI Symbol;Acc:MGI:109187]

Tnfrsf21	19.4854	12.3298	-0.66025	0.027421	ENSMUSG00000023915	tumor necrosis factor receptor superfamily, member 21 [Source:MGI Symbol;Acc:MGI:2151075]
Eid1	95.3184	60.1245	-0.6648	0.003238	ENSMUSG00000091337	EP300 interacting inhibitor of differentiation 1 [Source:MGI Symbol;Acc:MGI:1889651]
Sparc	464.698	292.689	-0.66693	0.020648	ENSMUSG00000018593	secreted acidic cysteine rich glycoprotein [Source:MGI Symbol;Acc:MGI:98373]
Vti1b	235.887	148.104	-0.67149	0.049457	ENSMUSG00000021124	vesicle transport through interaction with t-SNAREs 1B homolog [Source:MGI Symbol;Acc:MGI:1855688]
Rpl7a	78.0275	48.8241	-0.67639	0.036306	ENSMUSG00000062647	ribosomal protein L7A [Source:MGI Symbol;Acc:MGI:1353472]
Gapdh	978.936	612.471	-0.67657	0.018789	ENSMUSG00000057666	glyceraldehyde-3-phosphate dehydrogenase [Source:MGI Symbol;Acc:MGI:95640]
Tagln3	231.462	143.951	-0.68519	0.001289	ENSMUSG00000022658	transgelin 3 [Source:MGI Symbol;Acc:MGI:1926784]
Ypel3	85.6891	53.0646	-0.69136	0.016405	ENSMUSG00000042675	yippee-like 3 (Drosophila) [Source:MGI Symbol;Acc:MGI:1913340]
Ppp1r3c	18.1857	11.2202	-0.69671	0.043296	ENSMUSG00000067279	protein phosphatase 1, regulatory (inhibitor) subunit 3C [Source:MGI Symbol;Acc:MGI:1858229]
Fam32a	28.5826	17.6337	-0.69681	0.026388	ENSMUSG00000003039	family with sequence similarity 32, member A [Source:MGI Symbol;Acc:MGI:1915172]

Gm12328	1647.47	1014.84	-0.69899	0.01671	ENSMUSG00000081552	predicted gene 12328 [Source:MGI Symbol;Acc:MGI:3651322]
Cyc1	79.2877	48.7431	-0.7019	0.010233	ENSMUSG00000022551	cytochrome c-1 [Source:MGI Symbol;Acc:MGI:1913695]
Gprc5b	27.7689	16.9072	-0.71584	0.017547	ENSMUSG00000008734	G protein-coupled receptor, family C, group 5, member B [Source:MGI Symbol;Acc:MGI:1927596]
Cmas	54.5798	33.1457	-0.71954	0.007047	ENSMUSG00000030282	cytidine monophospho-N-acetylneuraminic acid synthetase [Source:MGI Symbol;Acc:MGI:1337124]
Cops4	56.7448	34.4071	-0.72178	0.015848	ENSMUSG00000035297	COP9 (constitutive photomorphogenic) homolog, subunit 4 (Arabidopsis thaliana) [Source:MGI Symbol;Acc:MGI:1349414]
Cab39	65.3814	39.5991	-0.72341	0.006599	ENSMUSG00000036707	calcium binding protein 39 [Source:MGI Symbol;Acc:MGI:107438]
Fbxo21	17.1586	10.3796	-0.72519	0.011476	ENSMUSG00000032898	F-box protein 21 [Source:MGI Symbol;Acc:MGI:1924223]
Tpi1	133.743	80.0706	-0.74012	0.048033	ENSMUSG00000023456	triosephosphate isomerase 1 [Source:MGI Symbol;Acc:MGI:98797]
Gap43	147.04	87.9553	-0.74137	0.000944	ENSMUSG00000047261	growth associated protein 43 [Source:MGI Symbol;Acc:MGI:95639]
Rpl26	215.837	127.76	-0.75651	0.033107	ENSMUSG00000060938	ribosomal protein L26 [Source:MGI Symbol;Acc:MGI:106022]
Gm12033	68.8541	40.4422	-0.76768	0.003862	ENSMUSG00000078965	predicted gene 12033 [Source:MGI Symbol;Acc:MGI:3650772]

Cct7	71.5359	41.9857	-0.76877	0.000268	ENSMUSG00000030007	chaperonin containing Tcp1, subunit 7 (eta) [Source:MGI Symbol;Acc:MGI:107184]
Reep2	20.6954	12.1394	-0.76961	0.022352	ENSMUSG00000038555	receptor accessory protein 2 [Source:MGI Symbol;Acc:MGI:2385070]
Sod1	740.558	433.616	-0.7722	9.21E-05	ENSMUSG00000022982	superoxide dismutase 1, soluble [Source:MGI Symbol;Acc:MGI:98351]
Ppp2cb	64.4952	37.738	-0.77317	0.003896	ENSMUSG00000009630	protein phosphatase 2 (formerly 2A), catalytic subunit, beta isoform [Source:MGI Symbol;Acc:MGI:1321161]
Rab7	111.372	65.1223	-0.77416	0.014408	ENSMUSG00000079477	RAB7, member RAS oncogene family [Source:MGI Symbol;Acc:MGI:105068]
Smarcd1	13.2763	7.69992	-0.78593	0.019919	ENSMUSG00000023018	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1 [Source:MGI Symbol;Acc:MGI:1933623]
Nfu1	60.971	35.2495	-0.79052	0.049002	ENSMUSG00000029993	NFU1 iron-sulfur cluster scaffold homolog (S. cerevisiae) [Source:MGI Symbol;Acc:MGI:1913290]
Qpct	22.3182	12.8668	-0.79457	0.034661	ENSMUSG00000024084	glutaminyl-peptide cyclotransferase (glutaminyl cyclase) [Source:MGI Symbol;Acc:MGI:1917786]
Gm5564	62.0448	35.5591	-0.80309	0.002233	ENSMUSG00000046440	predicted gene 5564 [Source:MGI Symbol;Acc:MGI:3645018]

1500011B03Rik, 2610524H06Rik, Gm20499	46.8134	26.8022	-0.80457	0.038156	#N/A	#N/A
Clptm1	26.5024	15.132	-0.80852	0.000751	ENSMUSG00000002981	cleft lip and palate associated transmembrane protein 1 [Source:MGI Symbol;Acc:MGI:1927155]
Ppme1	55.2024	31.4616	-0.81114	0.000266	ENSMUSG00000030718	protein phosphatase methylesterase 1 [Source:MGI Symbol;Acc:MGI:1919840]
2310003C23Rik	38.6234	21.9152	-0.81755	0.003057	ENSMUSG00000027573	RIKEN cDNA 2310003C23 gene [Source:MGI Symbol;Acc:MGI:1923675]
Hist3h2a,Trim17	24.823	14.0648	-0.81959	0.031079	#N/A	#N/A
Fau	185.993	105.243	-0.82153	0.03214	ENSMUSG00000038274	Finkel-Biskis-Reilly murine sarcoma virus (FBR-MuSV) ubiquitously expressed (fox derived) [Source:MGI Symbol;Acc:MGI:102547]
Nxph1	24.065	13.4797	-0.83615	0.022352	ENSMUSG00000046178	neurexophilin 1 [Source:MGI Symbol;Acc:MGI:107492]
4930547N16Rik	59.5699	33.3343	-0.83758	0.015518	ENSMUSG00000035365	RIKEN cDNA 4930547N16 gene [Source:MGI Symbol;Acc:MGI:1922567]
Uchl3	23.2096	12.9744	-0.83905	0.042928	ENSMUSG00000022111	ubiquitin carboxyl-terminal esterase L3 (ubiquitin thiolesterase) [Source:MGI Symbol;Acc:MGI:1355274]

Pacsin1	39.224	21.9135	-0.83991	0.010425	ENSMUSG00000040276	protein kinase C and casein kinase substrate in neurons 1 [Source:MGI Symbol;Acc:MGI:1345181]
Vapa	72.5406	40.2635	-0.84932	0.000422	ENSMUSG00000024091	vesicle-associated membrane protein, associated protein A [Source:MGI Symbol;Acc:MGI:1353561]
Cox6b1	363.621	201.591	-0.85101	0.000143	ENSMUSG00000036751	cytochrome c oxidase, subunit VIb polypeptide 1 [Source:MGI Symbol;Acc:MGI:107460]
Eif1a	29.7026	16.3065	-0.86514	0.043294	ENSMUSG00000057561	eukaryotic translation initiation factor 1A [Source:MGI Symbol;Acc:MGI:95298]
AC148327.1	263.986	144.619	-0.86821	0.000102	ENSMUSG00000096617	0
Gp1bb,Sept5	39.3636	21.5492	-0.86923	0.000285	#N/A	#N/A
Gm10166	77.759	42.2982	-0.87841	0.031079	ENSMUSG00000066543	predicted pseudogene 10166 [Source:MGI Symbol;Acc:MGI:3641818]
Ndfip1	825.726	448.459	-0.88069	6.53E-06	ENSMUSG00000024425	Nedd4 family interacting protein 1 [Source:MGI Symbol;Acc:MGI:1929601]
Pdhx	14.6715	7.95069	-0.88387	0.04676	ENSMUSG00000010914	pyruvate dehydrogenase complex, component X [Source:MGI Symbol;Acc:MGI:1351627]
Mag	22.8908	12.3754	-0.8873	0.002304	ENSMUSG00000036634	myelin-associated glycoprotein [Source:MGI Symbol;Acc:MGI:96912]
Trim35	14.7635	7.96348	-0.89056	0.010885	ENSMUSG00000022043	tripartite motif-containing 35 [Source:MGI Symbol;Acc:MGI:1914104]

Ranbp3	14.9197	8.04571	-0.89093	0.007331	ENSMUSG00000002372	RAN binding protein 3 [Source:MGI Symbol;Acc:MGI:1919060]
Hexim1	14.4224	7.72604	-0.90051	0.00328	ENSMUSG00000048878	hexamethylene bis-acetamide inducible 1 [Source:MGI Symbol;Acc:MGI:2385923]
Tusc2	32.9219	17.6292	-0.90108	0.002427	ENSMUSG00000010054	tumor suppressor candidate 2 [Source:MGI Symbol;Acc:MGI:1931086]
Sfr1	66.8826	35.549	-0.91182	5.04E-05	ENSMUSG00000025066	SWI5 dependent recombination repair 1 [Source:MGI Symbol;Acc:MGI:1915038]
Arglu1	15.83	8.34163	-0.92426	0.014834	ENSMUSG00000040459	arginine and glutamate rich 1 [Source:MGI Symbol;Acc:MGI:2442985]
Ctbp1	39.5104	20.8057	-0.92526	0.000148	ENSMUSG00000037373	C-terminal binding protein 1 [Source:MGI Symbol;Acc:MGI:1201685]
Wdr1	57.5128	30.0536	-0.93634	2.89E-06	ENSMUSG00000005103	WD repeat domain 1 [Source:MGI Symbol;Acc:MGI:1337100]
Fkbp1b	27.0899	14.0146	-0.95082	0.024464	ENSMUSG00000020635	FK506 binding protein 1b [Source:MGI Symbol;Acc:MGI:1336205]
Pcp4l1	21.0184	10.7915	-0.96175	0.009045	ENSMUSG00000038370	Purkinje cell protein 4-like 1 [Source:MGI Symbol;Acc:MGI:1913675]
Zfp62	93.4828	47.8476	-0.96625	0.028683	ENSMUSG00000046311	zinc finger protein 62 [Source:MGI Symbol;Acc:MGI:99662]
Gm15501	57.5952	29.2565	-0.97719	0.006599	ENSMUSG00000087412	predicted pseudogene 15501 [Source:MGI Symbol;Acc:MGI:3704296]

Rcl1	12.7669	6.41326	-0.99327	0.024435	ENSMUSG00000024785	RNA terminal phosphate cyclase-like 1 [Source:MGI Symbol;Acc:MGI:1913275]
Nudt3	198.284	99.5314	-0.99435	0.004073	ENSMUSG00000024213	nudix (nucleotide diphosphate linked moiety X)-type motif 3 [Source:MGI Symbol;Acc:MGI:1928484]
Oxct1	217.768	109.235	-0.99535	0.005753	ENSMUSG00000022186	3-oxoacid CoA transferase 1 [Source:MGI Symbol;Acc:MGI:1914291]
Aff2	28.8051	14.419	-0.99836	0.013952	ENSMUSG00000031189	AF4/FMR2 family, member 2 [Source:MGI Symbol;Acc:MGI:1202294]
Ndufv1	91.337	45.2477	-1.01336	0.015147	ENSMUSG00000037916	NADH dehydrogenase (ubiquinone) flavoprotein 1 [Source:MGI Symbol;Acc:MGI:107851]
6030419C18Rik	13.2118	6.50601	-1.02198	0.008062	ENSMUSG00000066607	RIKEN cDNA 6030419C18 gene [Source:MGI Symbol;Acc:MGI:2442108]
Vwa1	6.0212	2.96479	-1.02212	0.028683	ENSMUSG00000042116	von Willebrand factor A domain containing 1 [Source:MGI Symbol;Acc:MGI:2179729]
Amd-ps3	19.9873	9.82203	-1.02499	0.012455	ENSMUSG00000060096	S-adenosylmethionine decarboxylase, pseudogene 3 [Source:MGI Symbol;Acc:MGI:3704487]
Thbs2	2.74087	1.33689	-1.03576	0.037589	ENSMUSG00000023885	thrombospondin 2 [Source:MGI Symbol;Acc:MGI:98738]
2210016L21Rik	11.2131	5.43	-1.04616	0.038559	ENSMUSG00000029559	RIKEN cDNA 2210016L21 gene [Source:MGI Symbol;Acc:MGI:1919607]

Praf2	35.5601	17.1829	-1.04928	0.00053	ENSMUSG00000031149	PRA1 domain family 2 [Source:MGI Symbol;Acc:MGI:1859607]
Chchd10	126.298	60.9237	-1.05175	7.31E-06	ENSMUSG00000049422	coiled-coil-helix-coiled-coil-helix domain containing 10 [Source:MGI Symbol;Acc:MGI:2143558]
Gfap	60.2089	28.9459	-1.05662	2.43E-07	ENSMUSG00000020932	glial fibrillary acidic protein [Source:MGI Symbol;Acc:MGI:95697]
Uqcrh	415.303	198.519	-1.06489	7.71E-07	ENSMUSG00000063882	ubiquinol-cytochrome c reductase hinge protein [Source:MGI Symbol;Acc:MGI:1913826]
Hba-a1	218.668	104.258	-1.06859	7.10E-05	ENSMUSG00000069919	hemoglobin alpha, adult chain 1 [Source:MGI Symbol;Acc:MGI:96015]
Cldn11	58.393	27.7848	-1.0715	4.95E-06	ENSMUSG00000037625	claudin 11 [Source:MGI Symbol;Acc:MGI:106925]
Cnot8	31.1279	14.7672	-1.07581	0.004692	ENSMUSG00000020515	CCR4-NOT transcription complex, subunit 8 [Source:MGI Symbol;Acc:MGI:1916375]
2400001E08Rik	20.4639	9.5956	-1.09263	0.007955	ENSMUSG00000030842	RIKEN cDNA 2400001E08 gene [Source:MGI Symbol;Acc:MGI:1913758]
Serpina3n	24.1918	11.3298	-1.09439	7.53E-05	ENSMUSG00000021091	serine (or cysteine) peptidase inhibitor, clade A, member 3N [Source:MGI Symbol;Acc:MGI:105045]
Mrpl41	135.069	63.1819	-1.09612	3.39E-06	ENSMUSG00000036850	mitochondrial ribosomal protein L41 [Source:MGI Symbol;Acc:MGI:1333816]
Sccpdh	184.005	85.9126	-1.0988	0.005413	ENSMUSG00000038936	saccharopine dehydrogenase (putative) [Source:MGI Symbol;Acc:MGI:1924486]

Mif	140.608	65.2812	-1.10694	8.07E-05	ENSMUSG00000033307	macrophage migration inhibitory factor [Source:MGI Symbol;Acc:MGI:96982]
Prcc	7.23394	3.29701	-1.13362	0.030198	ENSMUSG00000004895	papillary renal cell carcinoma (translocation-associated) [Source:MGI Symbol;Acc:MGI:2137738]
Caly	129.765	58.7189	-1.144	1.48E-07	ENSMUSG00000025468	calcyon neuron-specific vesicular protein [Source:MGI Symbol;Acc:MGI:1915816]
Gm7292	95.4951	42.8587	-1.15584	1.04E-06	ENSMUSG00000072829	predicted gene 7292 [Source:MGI Symbol;Acc:MGI:3645786]
Tppp3	55.2844	24.7239	-1.16096	0.00012	ENSMUSG00000014846	tubulin polymerization-promoting protein family member 3 [Source:MGI Symbol;Acc:MGI:1915221]
Ankrd34a	7.10011	3.1725	-1.16222	0.026534	ENSMUSG00000049097	ankyrin repeat domain 34A [Source:MGI Symbol;Acc:MGI:3617846]
Tonsl,Vps28	124.811	55.5577	-1.16768	3.44E-07	#N/A	#N/A
Car2	78.2557	34.7263	-1.17217	2.67E-08	ENSMUSG00000027562	carbonic anhydrase 2 [Source:MGI Symbol;Acc:MGI:88269]
C030046I01Rik	15.655	6.93717	-1.1742	0.003488	ENSMUSG00000035781	RIKEN cDNA C030046I01 gene [Source:MGI Symbol;Acc:MGI:1924814]
Dyrk1a	54.3097	24.065	-1.17427	0.005408	ENSMUSG00000022897	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1a [Source:MGI Symbol;Acc:MGI:1330299]

Vtn	35.6159	15.3805	-1.21142	0.014254	ENSMUSG00000017344	vitronectin [Source:MGI Symbol;Acc:MGI:98940]
H3f3b	142.822	61.4378	-1.21702	1.38E-05	ENSMUSG00000016559	H3 histone, family 3B [Source:MGI Symbol;Acc:MGI:1101768]
Gm15920	181.262	77.8861	-1.21864	0.000678	ENSMUSG00000080893	predicted gene 15920 [Source:MGI Symbol;Acc:MGI:3801934]
Drg2	54.0491	22.8711	-1.24074	0.043722	ENSMUSG00000020537	developmentally regulated GTP binding protein 2 [Source:MGI Symbol;Acc:MGI:1342307]
Wibg	14.213	5.98364	-1.24812	0.013774	ENSMUSG00000064030	within bgcn homolog (Drosophila) [Source:MGI Symbol;Acc:MGI:1925678]
Cldn10	23.3534	9.80055	-1.2527	0.013783	ENSMUSG00000022132	claudin 10 [Source:MGI Symbol;Acc:MGI:1913101]
Syt5	51.0936	21.3181	-1.26106	7.78E-08	ENSMUSG00000004961	synaptotagmin V [Source:MGI Symbol;Acc:MGI:1926368]
Arf3	82.8923	34.5175	-1.26391	0.008894	ENSMUSG00000051853	ADP-ribosylation factor 3 [Source:MGI Symbol;Acc:MGI:99432]
Creld2	15.7127	6.50684	-1.2719	0.002814	ENSMUSG00000023272	cysteine-rich with EGF-like domains 2 [Source:MGI Symbol;Acc:MGI:1923987]
S100b	136.941	56.6325	-1.27385	5.36E-10	ENSMUSG00000033208	S100 protein, beta polypeptide, neural [Source:MGI Symbol;Acc:MGI:98217]
Gpr12	11.5666	4.73839	-1.2875	0.002037	ENSMUSG00000041468	G-protein coupled receptor 12 [Source:MGI Symbol;Acc:MGI:101909]
Eif4g2	189.493	76.9924	-1.29936	0.004499	ENSMUSG00000005610	eukaryotic translation initiation factor 4, gamma 2 [Source:MGI Symbol;Acc:MGI:109207]

Cadm3	44.4792	17.8522	-1.31703	0.034494	ENSMUSG00000005338	cell adhesion molecule 3 [Source:MGI Symbol;Acc:MGI:2137858]
1110008F13Rik	20.6297	8.26058	-1.32041	0.015518	ENSMUSG00000027637	RIKEN cDNA 1110008F13 gene [Source:MGI Symbol;Acc:MGI:1914638]
Sh3gl3	45.2781	18.0535	-1.32653	0.002643	ENSMUSG00000030638	SH3-domain GRB2-like 3 [Source:MGI Symbol;Acc:MGI:700011]
Mff	105.21	41.9486	-1.32657	0.01204	ENSMUSG00000026150	mitochondrial fission factor [Source:MGI Symbol;Acc:MGI:1922984]
Gm15427	26.2251	10.4101	-1.33295	0.020145	ENSMUSG00000081051	predicted pseudogene 15427 [Source:MGI Symbol;Acc:MGI:3642341]
SCARNA20	2278.39	895.726	-1.34688	0.031079	ENSMUSG00000089209	Small Cajal body specific RNA 20 [Source:RFAM;Acc:RF00601]
Necab2	15.7905	6.1062	-1.37071	0.020708	ENSMUSG00000031837	N-terminal EF-hand calcium binding protein 2 [Source:MGI Symbol;Acc:MGI:2152211]
Gfer	11.115	4.28969	-1.37356	0.003488	ENSMUSG00000040888	growth factor, erv1 (S. cerevisiae)-like (augmenter of liver regeneration) [Source:MGI Symbol;Acc:MGI:107757]
Mrpl48-ps	20.9756	7.86971	-1.41433	0.020708	ENSMUSG00000078480	mitochondrial ribosomal protein L48 pseudogene [Source:MGI Symbol;Acc:MGI:3050089]
Hadh	17.9244	6.69126	-1.42157	6.41E-05	ENSMUSG00000027984	hydroxyacyl-Coenzyme A dehydrogenase [Source:MGI Symbol;Acc:MGI:96009]

Tmem63b	27.7818	10.2526	-1.43815	0.019159	ENSMUSG00000036026	transmembrane protein 63b [Source:MGI Symbol;Acc:MGI:2387609]
Rtn4rl1	4.30095	1.56947	-1.45438	0.009276	ENSMUSG00000045287	reticulon 4 receptor-like 1 [Source:MGI Symbol;Acc:MGI:2661375]
Erh	85.9793	30.8921	-1.47675	0.000116	ENSMUSG00000021131	enhancer of rudimentary homolog (Drosophila) [Source:MGI Symbol;Acc:MGI:108089]
Ube2s	89.5018	31.8918	-1.48873	5.06E-09	ENSMUSG00000060860	ubiquitin-conjugating enzyme E2S [Source:MGI Symbol;Acc:MGI:1925141]
Gm9843	73.4298	25.9946	-1.49815	0.000102	ENSMUSG00000050299	predicted gene 9843 [Source:MGI Symbol;Acc:MGI:3708621]
Gpbp1l1	49.3353	16.5199	-1.57842	0.004499	ENSMUSG00000034042	GC-rich promoter binding protein 1-like 1 [Source:MGI Symbol;Acc:MGI:1924360]
Rbbp6	64.5126	21.0892	-1.61308	0.002423	ENSMUSG00000030779	retinoblastoma binding protein 6 [Source:MGI Symbol;Acc:MGI:894835]
Gm11353	26.8346	8.66772	-1.63037	0.00137	ENSMUSG00000060198	predicted gene 11353 [Source:MGI Symbol;Acc:MGI:3649794]
Mpnd	44.2849	14.1666	-1.64432	8.07E-05	ENSMUSG00000003199	MPN domain containing [Source:MGI Symbol;Acc:MGI:1915297]
Hba-a2	645.713	204.785	-1.65679	0	ENSMUSG00000069917	hemoglobin alpha, adult chain 2 [Source:MGI Symbol;Acc:MGI:96016]
Tsc22d4	38.7625	12.1985	-1.66796	0.012455	ENSMUSG00000029723	TSC22 domain family, member 4 [Source:MGI Symbol;Acc:MGI:1926079]

Hist1h1c	18.8825	5.93611	-1.66946	3.09E-05	ENSMUSG00000036181	histone cluster 1, H1c [Source:MGI Symbol;Acc:MGI:1931526]
Gm10108	56.4492	17.3242	-1.70417	0.028937	ENSMUSG00000062038	predicted pseudogene 10108 [Source:MGI Symbol;Acc:MGI:3642431]
Lrrc4b	35.533	10.675	-1.73493	3.38E-12	ENSMUSG00000047085	leucine rich repeat containing 4B [Source:MGI Symbol;Acc:MGI:3027390]
RNaseP_nuc	506.951	148.426	-1.7721	4.92E-12	ENSMUSG00000088874	Nuclear RNase P [Source:RFAM;Acc:RF00009]
Pvalb	96.614	27.0589	-1.83613	1.76E-13	ENSMUSG00000005716	parvalbumin [Source:MGI Symbol;Acc:MGI:97821]
Ppp1r14a	40.2998	10.9388	-1.88132	0.000233	ENSMUSG00000037166	protein phosphatase 1, regulatory (inhibitor) subunit 14A [Source:MGI Symbol;Acc:MGI:1931139]
Rpl17-ps5	313.336	84.7551	-1.88634	0	ENSMUSG00000081855	ribosomal protein L17, pseudogene 5 [Source:MGI Symbol;Acc:MGI:3704246]
Timm23	418.651	83.9938	-2.31739	1.08E-05	ENSMUSG00000013701	translocase of inner mitochondrial membrane 23 homolog (yeast) [Source:MGI Symbol;Acc:MGI:1858317]
U12	379.431	58.8483	-2.68877	0.01513	ENSMUSG00000088188	U12 minor spliceosomal RNA [Source:RFAM;Acc:RF00007]
Snora74a	2191.22	263.94	-3.05345	0	ENSMUSG00000065649	small nucleolar RNA, H/ACA box 74A [Source:MGI Symbol;Acc:MGI:3646921]
Gm10268	16.8519	1.80542	-3.2225	0.026534	ENSMUSG00000066475	predicted gene 10268 [Source:MGI Symbol;Acc:MGI:3642637]

U11	1263.25	119.118	-3.40668	0.000105	ENSMUSG00000077317	U11 spliceosomal RNA [Source:RFAM;Acc:RF00548]
AC150277.1	7465.48	582.269	-3.68048	0	ENSMUSG00000092819	0
Snora7a	1688.57	115.675	-3.86765	0.009075	ENSMUSG00000064563	small nucleolar RNA, H/ACA box 7A [Source:MGI Symbol;Acc:MGI:3819512]

Appendix Table 5. Significantly differentially expressed genes

Significantly differentially expressed genes identified after mapping of the Illumina sequencing reads via TopHat 2.0.4 to the mm10 build of the mouse genome. Differential expression testing was performed using Cufflinks 2.0.2. Genes are ordered from most upregulated in *Gnasxl*$^{m+/p-}$ mice (KO) compared to wildtype (WT) mice.

Published works

Publications

Holmes AP, Wood SH, Merry BJ, de Magalhães JP (2013) A-to-I RNA editing does not change with age in the healthy male rat brain. Biogerontology **14**: 395-400

Yu C, Li Y, Holmes A, Szafranski K, Faulkes CG, Coen CW, Buffenstein R, Platzer M, de Magalhães JP, Church GM (2011) RNA sequencing reveals differential expression of mitochondrial and oxidation reduction genes in the long-lived naked mole-rat when compared to mice. PLos One **6**: e26729

Posters

Holmes A, Wood S, Plagge A, de Magalhaes JP**. Deciphering the roles of Gnasxl by RNAseq analysis.** Ageing and Basic Bioscience, Babraham Institute, Cambridge. September 2012

Holmes A, de Magalhaes JP, Plagge A. **RNAseq and histological analyses of hypothalami of lean and hypermetabolic Gnasxl knockout mice reveals reduced levels of the astrocyte marker Gfap.** Programming Obesity: Central and Peripheral Contributors, University of Cambridge. 2013

References

Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA (2010) A map of human genome variation from population-scale sequencing. Nature **467**: 1061–73

Akbarian S, Smith MA, Jones EG (1995) Editing for an AMPA receptor subunit RNA in prefrontal cortex and striatum in Alzheimer's disease, Huntington's disease and schizophrenia. Brain Res **699**: 297–304

Alvarez-Bolado G, Paul FA, Blaess S (2012) Sonic hedgehog lineage in the mouse hypothalamus: from progenitor domains to hypothalamic regions. Neural Dev **7**: 4

Atanasov AG, Nashev LG, Schweizer RAS, Frick C, Odermatt A (2004) Hexose-6-phosphate dehydrogenase determines the reaction direction of 11beta-hydroxysteroid dehydrogenase type 1 as an oxoreductase. FEBS Lett **571**: 129–33

Athanasiadis A, Rich A, Maas S (2004) Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol **2**: e391

Balland E, Dam J, Langlet F, Caron E, Steculorum S, Messina A, Rasika S, Falluel-Morel A, Anouar Y, Dehouck B, et al (2014) Hypothalamic tanycytes are an ERK-gated conduit for leptin into the brain. Cell Metab **19**: 293–301

Bánhegyi G, Benedetti A, Fulceri R, Senesi S (2004) Cooperativity between 11beta-hydroxysteroid dehydrogenase type 1 and hexose-6-phosphate dehydrogenase in the lumen of the endoplasmic reticulum. J Biol Chem **279**: 27017–21

Bánhegyi G, Csala M, Benedetti A (2009) Hexose-6-phosphate dehydrogenase: linking endocrinology and metabolism in the endoplasmic reticulum. J Mol Endocrinol **42**: 283–9

Barbon A, Barlati S (2000) Genomic organization, proposed alternative splicing mechanisms, and RNA editing structure of GRIK1. Cytogenet Cell Genet **88**: 236–9

Barbon A, Vallini I, La Vie L, Marchina E, Barlati S (2003) Glutamate receptor RNA editing: a molecular analysis of GluR2, GluR5 and GluR6 in human brain tissues and in NT2 cells following in vitro neural differentiation. Mol Brain Res **117**: 168–178

Bass BL, Nishikura K, Keller W, Seeburg PH, Emeson RB, O'Connell MA, Samuel CE, Herbert A (1997) A standardized nomenclature for adenosine deaminases that act on RNA. RNA **3**: 947–9

Bélanger M, Allaman I, Magistretti PJ (2011) Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation. Cell Metab **14**: 724–38

Van den Berge SA, Middeldorp J, Zhang CE, Curtis MA, Leonard BW, Mastroeni D, Voorn P, van de Berg WDJ, Huitinga I, Hol EM (2010) Longterm quiescent cells in the aged

human subventricular neurogenic system specifically express GFAP-delta. Aging Cell **9**: 313–26

Bettscheider M, Murgatroyd C, Spengler D (2011) Simultaneous DNA and RNA isolation from brain punches for epigenetics. BMC Res Notes **4**: 314

Blacker D, Wilcox MA, Laird NM, Rodes L, Horvath SM, Go RC, Perry R, Watson B, Bassett SS, McInnis MG, et al (1998) Alpha-2 macroglobulin is genetically associated with Alzheimer disease. Nat Genet **19**: 357–60

Blow M, Futreal PA, Wooster R, Stratton MR (2004) A survey of RNA editing in human brain. Genome Res **14**: 2379–87

Boer K, Middeldorp J, Spliet WGM, Razavi F, van Rijen PC, Baayen JC, Hol EM, Aronica E (2010) Immunohistochemical characterization of the out-of frame splice variants GFAP Delta164/Deltaexon 6 in focal lesions associated with chronic epilepsy. Epilepsy Res **90**: 99–109

Bonfiglio JJ, Inda C, Refojo D, Holsboer F, Arzt E, Silberstein S (2011) The corticotropin-releasing hormone network and the hypothalamic-pituitary-adrenal axis: molecular and cellular mechanisms involved. Neuroendocrinology **94**: 12–20

Bordone L, Guarente L (2005) Calorie restriction, SIRT1 and metabolism: understanding longevity. Nat Rev Mol Cell Biol **6**: 298–305

Bouret SG, Simerly RB (2006) Developmental programming of hypothalamic feeding circuits. Clin Genet **70**: 295–301

Brusa R, Zimmermann F, Koh D-S, Feldmeyer D, Gass P, Seeburg PH, Sprengel R (1995) Early-Onset Epilepsy and Postnatal Lethality Associated with an Editing-Deficient GluR-B Allele in Mice. Science (80-) **270**: 1677–1680

Buckman LB, Thompson MM, Moreno HN, Ellacott KLJ (2013) Regional astrogliosis in the mouse hypothalamus in response to obesity. J Comp Neurol **521**: 1322–33

Buffenstein R (2005) The naked mole-rat: a new long-living model for human aging research. J Gerontol A Biol Sci Med Sci **60A**: 1369–1377

Buffenstein R (2008) Negligible senescence in the longest living rodent, the naked mole-rat: insights from a successfully aging species. J Comp Physiol B **178**: 439–45

Bujalska IJ, Draper N, Michailidou Z, Tomlinson JW, White PC, Chapman KE, Walker EA, Stewart PM (2005) Hexose-6-phosphate dehydrogenase confers oxo-reductase activity upon 11 beta-hydroxysteroid dehydrogenase type 1. J Mol Endocrinol **34**: 675–84

Burns CM, Chu H, Rueter SM, Hutchinson LK, Canton H, Sanders-Bush E, Emeson RB (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. Nature **387**: 303–8

Carlson NG, Howard J, Gahring LC, Rogers SW (2000) RNA editing (Q/R site) and flop/flip splicing of AMPA receptor transcripts in young and old brains. Neurobiol Aging **21**: 599–606

Cenci C, Barzotti R, Galeano F, Corbelli S, Rota R, Massimi L, Di Rocco C, O'Connell MA, Gallo A (2008) Down-regulation of RNA editing in pediatric astrocytomas: ADAR2 editing activity inhibits cell migration and proliferation. J Biol Chem **283**: 7251–60

Chen Z, Duan X (2011) Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. Methods Mol Biol **733**: 93–103

Chung K, Wallace J, Kim S-Y, Kalyanasundaram S, Andalman AS, Davidson TJ, Mirzabekov JJ, Zalocusky KA, Mattis J, Denisin AK, et al (2013) Structural and molecular interrogation of intact biological systems. Nature **497**: 332–7

Clutterbuck DR, Leroy A, O'Connell MA, Semple CAM (2005) A bioinformatic screen for novel A-I RNA editing sites reveals recoding editing in BC10. Bioinformatics **21**: 2590–5

Cook JL, Re RN, deHaro DL, Abadie JM, Peters M, Alam J (2008) The trafficking protein GABARAP binds to and enhances plasma membrane expression and function of the angiotensin II type 1 receptor. Circ Res **102**: 1539–47

Danecek P, Nellåker C, McIntyre RE, Buendia-Buendia JE, Bumpstead S, Ponting CP, Flint J, Durbin R, Keane TM, Adams DJ (2012) High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. Genome Biol **13**: 26

Deininger P (2011) Alu elements: know the SINEs. Genome Biol **12**: 236

Demetrius L (2005) Of mice and men. EMBO Rep **6**: S39–44

Di S, Malcher-Lopes R, Marcheselli VL, Bazan NG, Tasker JG (2005) Rapid glucocorticoid-mediated endocannabinoid release and opposing regulation of glutamate and gamma-aminobutyric acid inputs to hypothalamic magnocellular neurons. Endocrinology **146**: 4292–301

Di S, Maxson MM, Franco A, Tasker JG (2009) Glucocorticoids regulate glutamate and GABA synapse-specific retrograde transmission via divergent nongenomic signaling pathways. J Neurosci **29**: 393–401

Edmonson MN, Zhang J, Yan C, Finney RP, Meerzaman DM, Buetow KH (2011) Bambino: a variant detector and alignment viewer for next-generation sequencing data in the SAM/BAM format. Bioinformatics **27**: 865–6

Eggington JM, Greene T, Bass BL (2011) Predicting sites of ADAR editing in double-stranded RNA. Nat Commun **2**: 319

Eisenberg E, Adamsky K, Cohen L, Amariglio N, Hirshberg A, Rechavi G, Levanon EY (2005a) Identification of RNA editing sites in the SNP database. Nucleic Acids Res **33**: 4612–7

Eisenberg E, Nemzer S, Kinar Y, Sorek R, Rechavi G, Levanon EY (2005b) Is abundant A-to-I RNA editing primate-specific? Trends Genet **21**: 77–81

Ensterö M, Daniel C, Wahlstedt H, Major F, Ohman M (2009) Recognition and coupling of A-to-I edited sites are determined by the tertiary structure of the RNA. Nucleic Acids Res **37**: 6916–26

Farajollahi S, Maas S (2010) Molecular diversity through RNA editing: a balancing act. Trends Genet **26**: 221–30

Feng Y, Sansam CL, Singh M, Emeson RB (2006) Altered RNA editing in mice lacking ADAR2 autoregulation. Mol Cell Biol **26**: 480–8

Fernandez-Galaz MC, Martinez Muñoz R, Villanua MA, Garcia-Segura LM (1999) Diurnal oscillation in glial fibrillary acidic protein in a perisuprachiasmatic area and its relationship to the luteinizing hormone surge in the female rat. Neuroendocrinology **70**: 368–76

Freeman MR, Rowitch DH (2013) Evolving concepts of gliogenesis: a look way back and ahead to the next 25 years. Neuron **80**: 613–23

Galeano F, Leroy A, Rossetti C, Gromova I, Gautier P, Keegan LP, Massimi L, Di Rocco C, O'Connell MA, Gallo A (2010) Human BLCAP transcript: new editing events in normal and cancerous tissues. Int J Cancer **127**: 127–37

Garatachea N, Emanuele E, Calero M, Fuku N, Arai Y, Abe Y, Murakami H, Miyachi M, Yvert T, Verde Z, et al (2014) ApoE gene and exceptional longevity: Insights from three independent cohorts. Exp Gerontol **53**: 16–23

Garavelli L, Pedori S, Zanacca C, Caselli G, Loiodice A, Mantovani G, Ammenti A, Virdis R, Banchini G (2005) Albright's hereditary osteodystrophy (pseudohypoparathyroidism type Ia): clinical case with a novel mutation of GNAS1. Acta Biomed **76**: 45–8

Garcia-Segura LM, Lorenz B, DonCarlos LL (2008) The role of glia in the hypothalamus: implications for gonadal steroid feedback and reproductive neuroendocrine output. Reproduction **135**: 419–29

Geneviève D, Sanlaville D, Faivre L, Kottler M-L, Jambou M, Gosset P, Boustani-Samara D, Pinto G, Ozilou C, Abeguilé G, et al (2005) Paternal deletion of the GNAS imprinted locus (including Gnasxl) in two girls presenting with severe pre- and post-natal growth retardation and intractable feeding difficulties. Eur J Hum Genet **13**: 1033–9

Goldman BD, Goldman SL, Lanz T, Magaurin A, Maurice A (1999) Factors influencing metabolic rate in naked mole-rats (Heterocephalus glaber). Physiol Behav **66**: 447–59

Gommans WM, Tatalias NE, Sie CP, Dupuis D, Vendetti N, Smith L, Kaushal R, Maas S (2008) Screening of human SNP database identifies recoding sites of A-to-I RNA editing. RNA **14**: 2074–85

Groeneweg FL, Karst H, de Kloet ER, Joëls M (2011) Rapid non-genomic effects of corticosteroids and their role in the central stress response. J Endocrinol **209**: 153–67

Han X, Chen M, Wang F, Windrem M, Wang S, Shanz S, Xu Q, Oberheim NA, Bekar L, Betstadt S, et al (2013) Forebrain engraftment by human glial progenitor cells enhances synaptic plasticity and learning in adult mice. Cell Stem Cell **12**: 342–53

Higuchi M, Maas S, Single FN, Hartner J, Rozov A, Burnashev N, Feldmeyer D, Sprengel R, Seeburg PH (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature **406**: 78–81

Higuchi M, Single FN, Köhler M, Sommer B, Sprengel R, Seeburg PH (1993) RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency. Cell **75**: 1361–1370

Hillhouse EW, Grammatopoulos DK (2013) The Molecular Mechanisms Underlying the Regulation of the Biological Activity of Corticotropin-Releasing Hormone Receptors: Implications for Physiology and Pathophysiology.

Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem **83**: 8604–10

Hindson CM, Chevillet JR, Briggs HA, Gallichotte EN, Ruf IK, Hindson BJ, Vessella RL, Tewari M (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Methods **10**: 1003–5

Holmes AP, Wood SH, Merry BJ, de Magalhães JP (2013) A-to-I RNA editing does not change with age in the healthy male rat brain. Biogerontology **14**: 395–400

Hoopengardner B, Bhalla T, Staber C, Reenan R (2003) Nervous system targets of RNA editing identified by comparative genomics. Science (80-) **301**: 832–6

Horsch M, Seeburg PH, Adler T, Aguilar-Pimentel JA, Becker L, Calzada-Wack J, Garrett L, Götz A, Hans W, Higuchi M, et al (2011) Requirement of the RNA editing enzyme ADAR2 for normal physiology in mice. J Biol Chem **286**: 18614–18622

Hsuchou H, He Y, Kastin AJ, Tu H, Markadakis EN, Rogers RC, Fossier PB, Pan W (2009) Obesity induces functional astrocytic leptin receptors in hypothalamus. Brain **132**: 889–902

Hyder CL, Isoniemi KO, Torvaldson ES, Eriksson JE (2011) Insights into intermediate filament regulation from development to ageing. J Cell Sci **124**: 1363–72

Ingham PW, Placzek M (2006) Orchestrating ontogenesis: variations on a theme by sonic hedgehog. Nat Rev Genet **7**: 841–50

Jacobsen CT, Miller RH (2003) Control of astrocyte migration in the developing cerebral cortex. Dev Neurosci **25**: 207–16

Jepson JEC, Savva YA, Yokose C, Sugden AU, Sahin A, Reenan RA (2010) Engineered alterations in RNA editing modulate complex behavior in Drosophila: regulatory diversity of adenosine deaminase acting on RNA (ADAR) targets. J Biol Chem **286**: 8325–8337

Johansen K, Lykkeboe G, Weber RE, Maloiy GM (1976) Blood respiratory properties in the naked mole rat Heterocephalus glaber, a mammal of low body temperature. Respir Physiol **28**: 303–14

Kamphuis W, Mamber C, Moeton M, Kooijman L, Sluijs JA, Jansen AHP, Verveer M, de Groot LR, Smith VD, Rangarajan S, et al (2012) GFAP isoforms in adult mouse brain with a focus on neurogenic astrocytes and reactive astrogliosis in mouse models of Alzheimer disease. PLoS One **7**: e42823

Kamphuis W, Middeldorp J, Kooijman L, Sluijs JA, Kooi E-J, Moeton M, Freriks M, Mizee MR, Hol EM (2014) Glial fibrillary acidic protein isoform expression in plaque related astrogliosis in Alzheimer's disease. Neurobiol Aging **35**: 492–510

Kastner S, Voss T, Keuerleber S, Glöckel C, Freissmuth M, Sommergruber W (2012) Expression of G protein-coupled receptor 19 in human lung cancer cells is triggered by entry into S-phase and supports G(2)-M cell-cycle progression. Mol Cancer Res **10**: 1343–58

Kikuno R, Nagase T, Waki M, Ohara O (2002) HUGE: a database for human large proteins identified in the Kazusa cDNA sequencing project. Nucleic Acids Res **30**: 166–168

Kim DDY, Kim TTY, Walsh T, Kobayashi Y, Matise TC, Buyske S, Gabriel A (2004) Widespread RNA editing of embedded alu elements in the human transcriptome. Genome Res **14**: 1719–25

Kim EB, Fang X, Fushan AA, Huang Z, Lobanov A V, Han L, Marino SM, Sun X, Turanov AA, Yang P, et al (2011) Genome sequencing reveals insights into physiology and longevity of the naked mole rat. Nature **479**: 223–7

Kim JG, Suyama S, Koch M, Jin S, Argente-Arizon P, Argente J, Liu Z-W, Zimmer MR, Jeong JK, Szigeti-Buck K, et al (2014) Leptin signaling in astrocytes regulates hypothalamic neuronal circuits and feeding. Nat Neurosci **17**: 908–10

Kiran A, Baranov P V (2010) DARNED: a DAtabase of RNa EDiting in humans. Bioinformatics **26**: 1772–6

Klimek-Tomczak K, Mikula M, Dzwonek A, Paziewska A, Karczmarski J, Hennig E, Bujnicki JM, Bragoszewski P, Denisenko O, Bomsztyk K, et al (2006) Editing of hnRNP K protein mRNA in colorectal adenocarcinoma and surrounding mucosa. Br J Cancer **94**: 586–92

Klok MD, Jakobsdottir S, Drent ML (2007) The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. Obes Rev **8**: 21–34

Kohama SG, Goss JR, Finch CE, McNeill TH (1995) Increases of glial fibrillary acidic protein in the aging female mouse brain. Neurobiol Aging **16**: 59–67

Krechowec SO, Burton KL, Newlaczyl AU, Nunn N, Vlatković N, Plagge A (2012) Postnatal changes in the expression pattern of the imprinted signalling protein XLαs underlie the changing phenotype of deficient mice. PLoS One **7**: e29753

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol **10**: R25

Larson J, Park TJ (2009) Extreme hypoxia tolerance of naked mole-rat brain. Neuroreport **20**: 1634–7

Lee DA, Bedont JL, Pak T, Wang H, Song J, Miranda-Angulo A, Takiar V, Charubhumi V, Balordi F, Takebayashi H, et al (2012) Tanycytes of the hypothalamic median eminence form a diet-responsive neurogenic niche. Nat Neurosci **15**: 700–2

Lehmann KA, Bass BL (2000) Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. Biochemistry **39**: 12875–84

Levanon EY, Eisenberg E, Yelin R, Nemzer S, Hallegger M, Shemesh R, Fligelman ZY, Shoshan A, Pollock SR, Sztybel D, et al (2004) Systematic identification of abundant A-to-I editing sites in the human transcriptome. Nat Biotechnol **22**: 1001–5

Levanon EY, Hallegger M, Kinar Y, Shemesh R, Djinovic-Carugo K, Rechavi G, Jantsch MF, Eisenberg E (2005) Evolutionarily conserved human targets of adenosine to inosine RNA editing. Nucleic Acids Res **33**: 1162–8

Li C, Chen P, Palladino A, Narayan S, Russell LK, Sayed S, Xiong G, Chen J, Stokes D, Butt YM, et al (2010) Mechanism of hyperinsulinism in short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency involves activation of glutamate dehydrogenase. J Biol Chem **285**: 31806–18

Li JB, Levanon EY, Yoon J-K, Aach J, Xie B, Leproust E, Zhang K, Gao Y, Church GM (2009) Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. Science (80-) **324**: 1210–3

Liang S, Mele J, Wu Y, Buffenstein R, Hornsby PJ (2010) Resistance to experimental tumorigenesis in cells of a long-lived mammal, the naked mole-rat (Heterocephalus glaber). Aging Cell **9**: 626–35

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell **133**: 523–36

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M (2012) Comparison of next-generation sequencing systems. J Biomed Biotechnol **2012**: 251364

Liu Z, Turan S, Wehbi VL, Vilardaga J-P, Bastepe M (2011) Extra-long Gαs variant XLαs protein escapes activation-induced subcellular redistribution and is able to provide sustained signaling. J Biol Chem **286**: 38558–69

Lomeli H, Mosbacher J, Melcher T, Hoger T, Kuner T, Monyer H, Higuchi M, Bach A, Seeburg P (1994) Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. Science (80-) **266**: 1709–1713

Ma H, Li R, Zhang Z, Tong T (2004) mRNA level of alpha-2-macroglobulin as an aging biomarker of human fibroblasts in culture. Exp Gerontol **39**: 415–21

Maas S, Kawahara Y, Tamburro KM, Nishikura K (2006) A-to-I RNA editing and human disease. RNA Biol **3**: 1–9

Maas S, Rich A, Nishikura K (2003) A-to-I RNA editing: recent news and residual mysteries. J Biol Chem **278**: 1391–4

De Magalhães JP, Budovsky A, Lehmann G, Costa J, Li Y, Fraifeld V, Church GM (2009) The Human Ageing Genomic Resources: online databases and tools for biogerontologists. Aging Cell **8**: 65–72

De Magalhães JP, Costa J, Church GM (2007) An analysis of the relationship between metabolism, developmental schedules, and longevity using phylogenetic independent contrasts. J Gerontol A Biol Sci Med Sci **62**: 149–60

Mantovani G (2011) Clinical review: Pseudohypoparathyroidism: diagnosis and treatment. J Clin Endocrinol Metab **96**: 3020–30

Martinez HD, Jasavala RJ, Hinkson I, Fitzgerald LD, Trimmer JS, Kung H-J, Wright ME (2008) RNA editing of androgen receptor gene transcripts in prostate cancer cells. J Biol Chem **283**: 29938–49

Masoro EJ (2005) Overview of caloric restriction and ageing. Mech Ageing Dev **126**: 913–22

Mathews DH, Moss WN, Turner DH (2010) Folding and Finding RNA Secondary Structure. Cold Spring Harb Perspect Biol **2**: a003665

Merry BJ, Kirk AJ, Goyns MH (2008) Dietary lipoic acid supplementation can mimic or block the effect of dietary restriction on life span. Mech Ageing Dev **129**: 341–8

Micallef L, Rodgers P (2014) eulerAPE: drawing area-proportional 3-Venn diagrams using ellipses. PLoS One **9**: e101717

Montano M, Long K (2011) RNA surveillance-An emerging role for RNA regulatory networks in aging. Ageing Res Rev **10**: 216–24

Morabito M V, Emeson RB (2009) RNA editing as a therapeutic target for CNS disorders. Neuropsychopharmacology **34**: 246

Morse DP, Bass BL (1999) Long RNA hairpins that contain inosine are present in Caenorhabditis elegans poly(A)+ RNA. Proc Natl Acad Sci **96**: 6048–6053

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods **5**: 621–8

Mukaide M, Sugiyama M, Korenaga M, Murata K, Kanto T, Masaki N, Mizokami M (2014) High-throughput and sensitive next-generation droplet digital PCR assay for the quantitation of the hepatitis C virus mutation at core amino acid 70. J Virol Methods **207**: 169–77

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science **320**: 1344–9

Nawashiro H, Messing A, Azzam N, Brenner M (1998) Mice lacking GFAP are hypersensitive to traumatic cerebrospinal injury. Neuroreport **9**: 1691–6

Ng FS, Tangredi MM, Jackson FR (2011) Glial cells physiologically modulate clock neurons and circadian behavior in a calcium-dependent manner. Curr Biol **21**: 625–34

Nicholas A, de Magalhaes JP, Kraytsberg Y, Richfield EK, Levanon EY, Khrapko K (2010) Age-related gene-specific changes of A-to-I mRNA editing in the human brain. Mech Ageing Dev **131**: 445–7

Nishikura K (2010) Functions and Regulation of RNA Editing by ADAR Deaminases. Annu Rev Biochem **79**: 321–349

Niswender CM, Sanders-Bush E, Emeson RB (1998) Identification and characterization of RNA editing events within the 5-HT2C receptor. Ann N Y Acad Sci **861**: 38–48

Nunn N, Feetham CH, Martin J, Barrett-Jolley R, Plagge A (2013) Elevated blood pressure, heart rate and body temperature in mice lacking the XLαs protein of the Gnas locus is due to increased sympathetic tone. Exp Physiol **98**: 1432–45

O'Dowd BF, Nguyen T, Lynch KR, Kolakowski LF, Thompson M, Cheng R, Marchese A, Ng G, Heng HHQ, George SR (1996) A novel gene codes for a putative G protein-coupled receptor with an abundant expression in brain. FEBS Lett **394**: 325–329

O'Neill M, McPartlin J, Arthure K, Riedel S, McMillan N (2011) Comparison of the TLDA with the Nanodrop and the reference Qubit system. J Phys Conf Ser **307**: 012047

Ohlson J, Pedersen JS, Haussler D, Ohman M (2007) Editing modifies the GABA(A) receptor subunit alpha3. RNA **13**: 698–703

Palladino MJ, Keegan LP, O'Connell MA, Reenan RA (2000) A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. Cell **102**: 437–49

Park E, Williams B, Wold BJ, Mortazavi A (2012) RNA editing in the human ENCODE RNA-seq data. Genome Res **22**: 1626–33

Paz N, Levanon EY, Amariglio N, Heimberger AB, Ram Z, Constantini S, Barbash ZS, Adamsky K, Safran M, Hirschberg A, et al (2007) Altered adenosine-to-inosine RNA editing in human cancer. Genome Res **17**: 1586–95

Picardi E, Gallo A, Galeano F, Tomaselli S, Pesole G (2012) A novel computational strategy to identify A-to-I RNA editing sites by RNA-Seq data: de novo detection in human spinal cord tissue. PLoS One **7**: e44184

Picardi E, Horner DS, Chiara M, Schiavon R, Valle G, Pesole G (2010) Large-scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing. Nucleic Acids Res **38**: 4755–67

Plagge A, Gordon E, Dean W, Boiani R, Cinti S, Peters J, Kelsey G (2004) The imprinted signaling protein XL alpha s is required for postnatal adaptation to feeding. Nat Genet **36**: 818–26

Plagge A, Kelsey G, Germain-Lee EL (2008) Physiological functions of the imprinted Gnas locus and its protein variants Galpha(s) and XLalpha(s) in human and mouse. J Endocrinol **196**: 193–214

Porath HT, Carmi S, Levanon EY (2014) A genome-wide map of hyper-edited RNA reveals numerous new sites. Nat Commun **5**: 4726

Prosser RA, Edgar DM, Heller HC, Miller JD (1994) A possible glial role in the mammalian circadian clock. Brain Res **643**: 296–301

Qureshi IA, Mehler MF (2011) Non-coding RNA networks underlying cognitive disorders across the lifespan. Trends Mol Med **17**: 337–346

Ramaswami G, Zhang R, Piskol R, Keegan LP, Deng P, O'Connell MA, Li JB (2013) Identifying RNA editing sites using RNA sequencing data alone. Nat Methods **10**: 128–32

Reid IA (1992) Interactions between ANG II, sympathetic nervous system, and baroreceptor reflexes in regulation of blood pressure. Am J Physiol **262**: E763–78

Richard N, Molin A, Coudray N, Rault-Guillaume P, Jüppner H, Kottler M-L (2013) Paternal GNAS mutations lead to severe intrauterine growth retardation (IUGR) and provide evidence for a role of XLαs in fetal development. J Clin Endocrinol Metab **98**: E1549–56

Riedmann EM, Schopoff S, Hartner JC, Jantsch MF (2008) Specificity of ADAR-mediated RNA editing in newly identified targets. RNA **14**: 1110–8

Rueter SM, Dawson TR, Emeson RB (1999) Regulation of alternative splicing by RNA editing. Nature **399**: 75–80

Sauvageot C (2002) Molecular mechanisms controlling cortical gliogenesis. Curr Opin Neurobiol **12**: 244–249

Schaub M, Keller W (2002) RNA editing by adenosine deaminases generates RNA and protein diversity. Biochimie **84**: 791–803

Schulz N, Himmelbauer H, Rath M, van Weeghel M, Houten S, Kulik W, Suhre K, Scherneck S, Vogel H, Kluge R, et al (2011) Role of medium- and short-chain L-3-hydroxyacyl-CoA dehydrogenase in the regulation of body weight and thermogenesis. Endocrinology **152**: 4641–51

Sebastiani P, Montano M, Puca A, Solovieff N, Kojima T, Wang MC, Melista E, Meltzer M, Fischer SEJ, Andersen S, et al (2009) RNA editing genes associated with extreme old age in humans and with lifespan in C. elegans. PLoS One **4**: e8210

Seluanov A, Hine C, Azpurua J, Feigenson M, Bozzella M, Mao Z, Catania KC, Gorbunova V (2009) Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. Proc Natl Acad Sci **106**: 19352–7

Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker H V, Xu W, Richards DR, McDonald-Smith GP, Gao H, Hennessy L, et al (2013) Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc Natl Acad Sci U S A. doi: 10.1073/pnas.1222878110

Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, Delaney A, Gelmon K, Guliany R, Senz J, et al (2009) Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature **461**: 809–13

Shen Y, Fu W-Y, Cheng EYL, Fu AKY, Ip NY (2013) Melanocortin-4 receptor regulates hippocampal synaptic plasticity through a protein kinase A-dependent mechanism. J Neurosci **33**: 464–72

Shestakova A, Zolov S, Lupashin V (2006) COG complex-mediated recycling of Golgi glycosyltransferases is essential for normal protein glycosylation. Traffic **7**: 191–204

Sie CP, Maas S (2009) Conserved recoding RNA editing of vertebrate C1q-related factor C1QL1. FEBS Lett **583**: 1171–4

Simmons M, Meador-Woodruff JH, Sodhi MS (2010) Increased cortical expression of an RNA editing enzyme occurs in major depressive suicide victims. Neuroreport **21**: 993–7

Singh M, Kesterson RA, Jacobs MM, Joers JM, Gore JC, Emeson RB (2007) Hyperphagia-mediated obesity in transgenic mice misexpressing the RNA-editing enzyme ADAR2. J Biol Chem **282**: 22448–59

Sofroniew M V (2009) Molecular dissection of reactive astrogliosis and glial scar formation. Trends Neurosci **32**: 638–47

Solga AC, Pong WW, Walker J, Wylie T, Magrini V, Apicelli AJ, Griffith M, Griffith OL, Kohsaka S, Wu GF, et al (2015) RNA-sequencing reveals oligodendrocyte and neuronal transcripts in microglia relevant to central nervous system disease. Glia **63**: 531–48

Speakman JR, Mitchell SE (2011) Caloric restriction. Mol Aspects Med **32**: 159–221

Sreedharan S, Almén MS, Carlini VP, Haitina T, Stephansson O, Sommer WH, Heilig M, de Barioglio SR, Fredriksson R, Schiöth HB (2011) The G protein coupled receptor Gpr153 shares common evolutionary origin with Gpr162 and is highly expressed in central regions including the thalamus, cerebellum and the arcuate nucleus. FEBS J **278**: 4881–94

Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D, Costa J, Fraifeld VE, de Magalhães JP (2013) Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. Nucleic Acids Res **41**: D1027–33

Tarailo-Graovac M, Chen N (2009) Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics **Chapter 4**: Unit 4.10

Tian X, Azpurua J, Hine C, Vaidya A, Myakishev-Rempel M, Ablaeva J, Mao Z, Nevo E, Gorbunova V, Seluanov A (2013) High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. Nature **499**: 346–9

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25**: 1105–11

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol **28**: 511–515

Weinstein LS, Xie T, Qasem A, Wang J, Chen M (2010) The role of GNAS and other imprinted genes in the development of obesity. Int J Obes (Lond) **34**: 6–17

Wood SH, Craig T, Li Y, Merry B, Magalhães JP (2013) Whole transcriptome sequencing of the aging rat brain reveals dynamic RNA changes in the dark matter of the genome. Age (Omaha) **35**: 763–76

Xie T, Plagge A, Gavrilova O, Pack S, Jou W, Lai EW, Frontera M, Kelsey G, Weinstein LS (2006) The alternative stimulatory G protein alpha-subunit XLalphas is a critical regulator of energy and glucose metabolism and sympathetic nerve activity in adult mice. J Biol Chem **281**: 18989–99

Xu K, Malouf AT, Messing A, Silver J (1999) Glial fibrillary acidic protein is necessary for mature astrocytes to react to beta-amyloid. Glia **25**: 390–403

Yang L, Zhao L, Gan Z, He Z, Xu J, Gao X, Wang X, Han W, Chen L, Xu T, et al (2010) Deficiency in RNA editing enzyme ADAR2 impairs regulated exocytosis. FASEB J **24**: 3720–32

Yu C, Li Y, Holmes A, Szafranski K, Faulkes CG, Coen CW, Buffenstein R, Platzer M, de Magalhães JP, Church GM (2011) RNA sequencing reveals differential expression of mitochondrial and oxidation reduction genes in the long-lived naked mole-rat when compared to mice. PLoS One **6**: e26729

Yu S, Gavrilova O, Chen H, Lee R, Liu J, Pacak K, Parlow AF, Quon MJ, Reitman ML, Weinstein LS (2000) Paternal versus maternal transmission of a stimulatory G-protein alpha subunit knockout produces opposite effects on energy metabolism. J Clin Invest **105**: 615–23

Yu X, Sun S (2013) Comparing a few SNP calling algorithms using low-coverage sequencing data. BMC Bioinformatics **14**: 274

Zaranek AW, Levanon EY, Zecharia T, Clegg T, Church GM (2010) A Survey of Genomic Traces Reveals a Common Sequencing Error, RNA Editing, and DNA Editing. PLoS Genet **6**: e1000954

Zheng-Bradley X, Rung J, Parkinson H, Brazma A (2010) Large scale comparison of global gene expression patterns in human and mouse. Genome Biol **11**: R124

Zhu H, Urban DJ, Blashka J, McPheeters MT, Kroeze WK, Mieczkowski P, Overholser JC, Jurjus GJ, Dieter L, Mahajan GJ, et al (2012) Quantitative analysis of focused a-to-I RNA editing sites by ultra-high-throughput sequencing in psychiatric disorders. PLoS One **7**: e43227